

(TM)

Distribution rights by Oxford Molecular Ltd

- n.a. database search, using Smith-Waterman algorithm

MASSPAR time 3146.28 seconds

not generated.

TTAGACCGGAGCCTTGTGCG.....GTGAGACATGTTTCTTTT

## Cap 6

Dbase 0; Query 0

43/396 seqs, 834342348 bases X 2

**Listing first 45 summaries**

genbank1.06

25:gb\_sy 26:gb\_un 27:gb\_vl

Mean 12.226; Variance 7.675; scale 1.593

lived by analysis of the total score distribution.

## SUMMARIES

Accession	Length	DB	ID	Description	Prod. No.
AF038421	2560	22	Homo sapiens GPR-11like	0.00e+00	
HS095847	2175	22	Human GDNF receptor al	0.00e+00	
AF042080	1619	22	Homo sapiens glial cel	0.00e+00	
HS097144	1707	22	Homo sapiens RET ligand	0.00e+00	
RNU97144	3616	23	Rattus norvegicus RET	0.00e+00	
RNU55486	2138	23	Rattus norvegicus GDNF	0.00e+00	
AF014117	2549	23	Mus musculus GDNF rece	0.00e+00	
AF015172	1651	23	Mus musculus GDNF rece	0.00e+00	
AB000800	1415	23	Mus musculus GPR-11like	0.00e+00	
RNU207	1392	23	Rattus norvegicus mRNA	0.00e+00	
CGU90541	3037	17	Gallus gallus GDNF rec	0.00e+00	
HSGFR1G11	832	22	Homo sapiens GPR-11like	2.125e-29	
AF012811	777	22	Mus musculus GDNF rece	2.35e-29	
HSGFR1G06	596	22	Homo sapiens GPR-11like	3.42e-19	
HS097144	385	22	Homo sapiens GPR-11like	7.33e-16	

16	296	11.7	497	22	HSCFR1G03	Homo sapiens GPR-1like	1.57e-163
17	291	11.5	559	22	HSCFR1G02	Homo sapiens GPR-1like	3.29e-165
18	214	8.5	1392	23	AF002701	Mus musculus GDNF fam1	2.11e-112
19	214	8.5	1395	22	HS093703	Human glial cell line-	2.11e-112
20	214	8.5	1536	22	AF002700	Homo sapiens GDNF fam1	2.11e-112
21	210	8.5	2906	22	HS097145	Homo sapiens RET ligand	2.11e-112
22	210	8.3	2787	23	RN097143	Rattus norvegicus RET	8.71e-110
23	208	8.2	1395	23	AF005262	Rattus norvegicus glia	1.77e-108
24	183	7.2	432	22	HSCFR1G09	Homo sapiens GPR-1like	3.20e-92
25	166	6.6	2933	17	GG090542	Gallus gallus neuritin	3.15e-81
26	138	5.5	397	22	HSCFR1G08	Homo sapiens GPR-1like	2.47e-53
27	114	4.5	456	22	HSCFR1G07	Homo sapiens GPR-1like	2.95e-48
28	96	3.8	385	18	AA3902	Sequence 34 from patent	3.57e-37
29	85	3.4	363	22	HSCFR1G04	Homo sapiens GPR-1like	1.56e-30
30	30	2.5	7218	18	166494	Sequence 14 from patent	1.08e-17
31	58	2.3	256	22	HSCFR1G10	Homo sapiens GPR-1like	6.87e-15
32	31	2.0	7218	18	166494	Sequence 14 from patent	1.58e-10
33	33	1.3	215	18	128278	Sequence 5 from patent	4.93e-02
34	33	1.3	215	18	128278	Sequence 5 from patent	4.93e-02
35	33	1.3	1203	22	AF051167	Homo sapiens GDNF fam1	1.40e-01
36	37	1.3	4554	15	PFARPS2	P.falsiparum mRNA for	1.40e-01
37	34	1.3	10772	15	AF012089	Drosophila melanogaster	1.72e-02
38	33	1.3	10772	15	AF012089	Drosophila melanogaster	4.93e-02
39	30	1.2	1996	20	KC053121	Komodo caudata nucleomo	1.06e+00
40	40	1.2	7116	15	PFARPS1	Plasmodium falciparum	1.06e+00
41	42	1.2	137765	22	HS435P10	Human DNA sequence fto	1.06e+00
42	41	1.1	598	18	EO2043	5' untranslated region	2.85e+00
43	43	1.1	1021	27	ORCPV160	Comox virus (CPV) 160	2.85e+00
44	29	1.1	4654	27	CPVAT1	Comox virus A-type 1n	2.85e+00
45	29	1.1	4892	22	H05S01RNA	Homo sapiens sulfonyl in	2.85e+00

## ALIGNMENTS

TITLE		PRI	10-APR-1998
JOURNAL REFERENCE			
JOURNALS	Homo sapiens GPR1-linked anchor protein (GFR1) mRNA, complete cds.		
AUTHORS	Angrist,M., Jling,S., Bolk,S., Bentley,K., Nallasamy,S., Halushka,M., Fox,G.M. and Chakravarti,A. Human GPR1: Cloning, mapping, genomic structure, and evaluation as a candidate gene for Hirschsprung disease susceptibility		
CDS	Genomics 48; 354-362 (1998) 2.(bases 1 to 2560)		
FEATURES	Angrist,M., Jling,S., Bolk,S., Bentley,K., Nallasamy,S., Halushka,M., Fox,G.M. and Chakravarti,A.  Direct Submission Submitted (12-DEC-1997) Genetics, Case Western Reserve, 2109 Adelbert Road, Cleveland, OH 44106-4955, USA  location/Qualifiers 1..2560 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="10" /map="10q25" <1..>2560 /gene="GFR1" 557..1954 /gene="GFR1"		
gene	/note="similar to GDNF Receptor alpha; part of functional ligand for RET receptor tyrosine kinase" /codon_start=1 /product="GPI-linked anchor protein" /db_xref="pid:g2921533" translation "MFLATVLPALPDLLLSAEGSGDRIDCVKASQCKEKSCSTKRITLNCAGKETNSLASGLAEADDECRSNAAEAKRKSLYCNCKRRMKERKCCLIRI		





826 CTTTGACAAGGTCGGGGCAAGCAGACGATGGAATGCTCTCTGCTCTGCGGGACAT 885  
1136 CTTTGACAAGGTCGGGGCAAGCAGACGATGGAATGCTCTCTGCTCTGCGGGACAT 1195  
886 CGCCTGCACAGAGCGGAGGCGACAGACCATCTGCTGCTGCTCTATGAAGAGGGA 945  
1196 CGCCTGCACAGAGCGGAGGCGACAGACCATCTGCTGCTGCTCTATGAAGAGGGA 1255  
946 GAAGCCCAACCTGTTGAATTTGCAGAGACCTCTGCAAGCAATTTACATCTGCATCTCG 1005  
1236 GAAGCCCAACCTGTTGAATTTGCAGAGACCTCTGCAAGCAATTTACATCTGCATCTCG 1315  
1006 CTTTGCGGATTTTTTACCACTGCGCAGCAGAGTCAAGGCTGTCAGCAGCGCTGTAA 1065  
1316 CTTTGCGGATTTTTTACCACTGCGCAGCAGAGTCAAGGCTGTCAGCAGCGCTGTAA 1375  
1066 GGAACCTACGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125  
1376 GGAACCTACGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1435  
1126 CAACCTACAGACTCCAGTACGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1185  
1436 CAACCTACAGACTCCAGTACGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1495  
1186 GAAGCAGCTAGAGAGTGTGTAATTTCTTGAATTTCTTCAAGAGCAATACATGCTTAA 1245  
1496 GAAGCAGCTAGAGAGTGTGTAATTTCTTGAATTTCTTCAAGAGCAATACATGCTTAA 1555  
1246 AATGCAATTTCAAGCTTTGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1305  
1556 AATGCAATTTCAAGCTTTGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1615  
1306 AGTACAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1365  
1616 AGTACAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1675  
1366 GCGCAGAGGCTGTGAAATGAAATTTCCCACTGATGTTTGGCAGCGGTGTCGAAATTTACA 1425  
1676 GCGCAGAGGCTGTGAAATGAAATTTCCCACTGATGTTTGGCAGCGGTGTCGAAATTTACA 1735  
1426 GCGCAGAGGCTGTGAAATGAAATTTCCCACTGATGTTTGGCAGCGGTGTCGAAATTTACA 1485  
1736 GCGCAGAGGCTGTGAAATGAAATTTCCCACTGATGTTTGGCAGCGGTGTCGAAATTTACA 1795  
1486 TTATGAAAAAGAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1545  
1796 TTATGAAAAAGAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1855  
1546 TCACAAGCTGTGCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1605  
1856 TCACAAGCTGTGCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1915  
1606 ATCTTTAACAAGAACATCATCTAGCTGATTAATAAATAACAATATGAGCATGTAAGAAC 1665  
1916 ATCTTTAACAAGAACATCATCTAGCTGATTAATAAATAACAATATGAGCATGTAAGAAC 1975  
1666 AAAAAACAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1725  
1976 AAAAAACAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2035  
1726 AGTGTAGAAACAGTTCATCTCACTGAAACATTTTTTTTTTTTCCCTTTTAAAAAGCTTC 1785  
2036 AGTGTAGAAACAGTTCATCTCACTGAAACATTTTTTTTTTTTCCCTTTTAAAAAGCTTC 2095  
1786 TTGTGATCTTTGGGGCTTCTGTGAAAACTGTATGACAGTGTCTCATCCAACTCAGAAG 1845  
2096 TTGTGATCTTTGGGGCTTCTGTGAAAACTGTATGACAGTGTCTCATCCAACTCAGAAG 2155  
1846 GCTTGGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1905  
2156 GCTTGGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2215  
1906 GCGNTGTGTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1965

2216 GGGCTGTGTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2275  
1966 TCNTCATGATCATGATNATGATTTTAAACAGTTTACTTGTGCTTCTTCCATNANAA 2025  
2276 NNN 2335  
2026 NGAGTTATATTTCCANGCTACCTCCATCTCCTTAAAGACATGATGATTTCAATGAT 2085  
2336 GGAGTTATATTTCTAAGTATGCTCCATATCTCTTAAAGACATGATGATGATGATGATGAT 2395  
2086 ATTAATTTCACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2145  
2396 ATTAATTTCACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2455  
2146 NTGGGCTTTGTGAAACAGCTGTGATGAT 2175  
2456 GTGGGCTTTGTGAAACAGCTGTGATGAT 2485  
RESULT 3  
LOCUS AF042080 1619 bp mRNA PRI 22-JAN-1998  
DEFINITION Homo sapiens glial cell line-derived neurotrophic factor receptor  
ACCESSION alpha (GFRAL) mRNA, complete cds.  
NID AF042080  
KEYWORDS g2801556  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homindae; Homo.  
REFERENCE 1 (bases 1 to 1619)  
AUTHORS Shefelbine, S.E., Khourana, S., Schultz, P.N., Huang, E., Thobe, N.,  
Hu, Z.-J., Fox, G.M., Jing, S., Cole, G.J., and Gagel, R.F.  
TITLE Mutational analysis of the GDNF/RET-GDNFR signaling complex in a  
kindred with vesicoureteral reflux  
HUM. GENET. (1998) In press  
JOURNAL Shefelbine, S.E., Khourana, S., Schultz, P.N., Huang, E., Thobe, N.,  
Hu, Z.-J., Fox, G.M., Jing, S., Cole, G.J., and Gagel, R.F.  
AUTHORS Direct Submission  
TITLE Submitted (08-JAN-1998) Endocrinology-Box 15, M.D. Anderson Cancer  
JOURNAL Center, 1515 Holcombe Blvd, Houston, TX 77030, USA  
FEATURES  
source  
1..1619  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q25-q26"  
/tissue-type="medullary thyroid carcinoma tumor"  
1..1619  
/note="GDNFRa"  
/gene="GFRAL"  
111..1508  
/note="GPI-linked receptor"  
/codon\_start=1  
/product="glial cell line-derived neurotrophic factor  
receptor alpha"  
/db\_xref="PID:g2801557"  
/translation="MFLATLYFALPLDLLSAEVSGGRDLCVKASDQCEKQSGST  
KYRLROCVAGKEDNFSLASGLAEKDEKRSAMALOKSLYCRKRGKREKRLRI  
YMSXYSIOGNDLSDSPVPSRSLDIFRRVPTSLDVOQVEHLPKGNCLDAKRA  
CNLDIDCKTRATYITPCTTSVNDVCRNRKCKALRQFDKYPAAKSHYGLCSCRD  
IACETERRQTIIVPCSYIEERKPCNLNLDQSCNTYICRSLADFTNQCPEBSVSS  
CLKENYADCLAYSGLIGTVMPTNYIDSSLSYAPMCDNSGNDSECEKLFNFRD  
NFCIKNALIOFGNSDVTWQPAFPVQTTATTTALRYNKRPLGFGSENSEIPTHLV  
PCANLOAOKLRKNSVGNHLICISNGYKEKGEGLGASHITTKMAAPPSGGLPLVLV  
VYVRLSLSLSTETS"  
misc.feature  
529..543  
/gene="GFRAL"  
/note="found to be alternative exon after partial analysis  
of genomic sequence"







DEFINITION Rattus norvegicus RET ligand 1 (RET1) mRNA, complete cds.  
ACCESSION U97142  
NID 92282021  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;  
Murinae; Rattus.  
REFERENCE 1 (bases 1 to 3616)  
Santicola, M., Hession, C. A., Worley, D. S., Carmilho, P., Ehrenfels, C.,  
Walus, L., Robinson, S., Jaworski, G., Wei, H., Tizard, R., Whitley, A.,  
Pepinsky, R. B. and Cate, R. L.  
Glial cell line-derived neurotrophic factor-dependent RET  
activation can be mediated by two different cell-surface accessory  
proteins  
Proc. Natl. Acad. Sci. U.S.A. 94 (12), 6238-6243 (1997)  
JOURNAL 97322356  
MEDLINE 2 (bases 1 to 3616)  
REFERENCE Santicola, M., Hession, C. A., Worley, D. S., Carmilho, P., Ehrenfels, C.,  
Walus, L., Robinson, S., Jaworski, G., Wei, H., Tizard, R., Whitley, A.,  
Pepinsky, R. B. and Cate, R. L.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-1997) Molecular Genetics, BIOGEN, 14 Cambridge  
Center, Cambridge, MA 02142, USA  
FEATURES  
source  
1. 3616  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/dev\_stage="embryonic day 18"  
/tissue\_type="kidney"  
1. 3616  
/gene="RET1"  
257. 1663  
/note="PI-linked cell-surface accessory protein; GDNF  
receptor alpha"  
/codon\_start=1  
/product="RET ligand 1"  
/db\_xref="PID:92282022"  
/translation="MFLATLYFALPLDLMSAEVSGGRDLCVAKASDQCKEKSCT  
KYSRLROCVAKENFSLTSGLEAKDECSRAMEALIKOKSLNCRKRMKREKCLRI  
YMSWOSLOGNDLEDSPLEPNVNSRLDFRAVEFISFOOVEHISGNCLDAKA  
CNLDCTKKYSAYITPCTTSMNENVCRRCHALRFENKVAKHSYGLFGSCR  
IACERBKQITVPCSYERERPNCLSLQDSCKNIRSLADEFTTCQESSVSN  
CLKENYACCLAIYGLIGTVPNTVDSLSVAWCCSNKSGNDELCKLFNFKD  
NTCLNIAQAFNGNSDVTMOPAPVQTTATTTAFVKNKPLAPASENIEPTHLV  
PPCANLQAKLKSNSGSTHCLSDSDGKDLGASASHITTKMAAPSCSLSPV  
LMLTALALLSVSLAETS"

BASE COUNT 876 a 955 c 921 g 864 t  
ORIGIN

Query Match 49.2%; Score 1243; DB 23; Length 3616;  
Best Local Similarity 83.9%; Pred. No. 0.00e+00;  
Matches 1657; Conservative 0; Mismatches 297; Indels 20; Gaps 13;

Db 8 CAGGTTGGGTGCGGAACCTGAACCCCTGAAGCGGGTCCGCTCCCGCCCGCGCCGCC 67  
Qy 291 CAGGTTGGGTGCGGACTGAACCCCTGAAGCGGAACCGGCTCCCGCCCGCGC -CATGCC 348  
Db 68 GGATCTGAGTCGCTGGCGCGGGGCGGCGAGAGCGAGCGGAGTCTGCTCACCCCTG 127  
Qy 349 GGAGCTGAGTCCCGCGCGCGGGTGG -CTGCTCCAGAACCCGAGAGTTTCTCTTCACTGG 407  
Db 128 ATGAGAGCTGAACCTTGAAGTGCGGAGAGAGCGAGTCCCGCGGATCGCTCAGAGCTGA 187  
Qy 408 ATGAGAGCTGAACCTTGGGGGCGGAGAGAGAGAGCTGTCGGGGATCGCTCAGAGCTGA 467  
Db 188 GCTCTCTCCCGAGAGCGGGCGCGGCTTGGATTTT---GGGGGGGCGGGAGACAGCTG 244  
Qy 468 GCTCTCTCCCGAGAGCGGGCGCGGCTTGGATTTTGGGGGGGCGGGAGACAGCTG 527

Db 245 CGCGCGGACCAATGTTCTAGACACTGTGACTTGGCGGTGCCACTCTCGGATTTCTG 304  
Qy 528 CGCGCGGACCAATGTTCTAGACACTGTGACTTGGCGGTGCCACTCTCGGATTTCTG 587  
Db 305 ATGTCGCCCGAGGTGAGTGTGAGAGACCGTGTGACACTGTGAAAAGCCAGAGTCAAGTGC 364  
Qy 588 CTGTGCGCGGAGGTGAGGTGAG 647  
Db 365 CTGAAG 424  
Qy 648 CTGAAG 707  
Db 425 GAAACCAACTTCAAGCTGACATCCGCGCTGAGGCGCAAGATGATGATGATGATGATGATG 484  
Qy 708 GAGACCAACTTCAAGCTGACATCCGCGCTGAGGCGCAAGATGATGATGATGATGATGATG 767  
Db 485 GAGGCTTGAAG 544  
Qy 768 GAGGCTTGAAG 827  
Db 545 AAGATGTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604  
Qy 828 AAGATGTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887  
Db 605 GAGATTTCCCGTATGAGCCGCTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664  
Qy 888 GAGATTTCCCGTATGAGCCGCTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 947  
Db 665 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 724  
Qy 948 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1007  
Db 725 GCAGCGAGGCTGCAACCTGAG 784  
Qy 1008 GCAGCGAGGCTGCAACCTGAG 1067  
Db 785 CCTGACACACAGAGATGCTCAAG 844  
Qy 1068 CCTGACACACAGAGATGCTCAAG 1127  
Db 845 AGCAGTTCTTGCAG 904  
Qy 1128 CGGCGAGTTCTTGCAG 1187  
Db 905 CGGCGAGTTCTTGCAG 964  
Qy 1188 CGGCGAGTTCTTGCAG 1247  
Db 965 GAAACAG 1024  
Qy 1248 GAAACAG 1307  
Db 1025 AGATCTGCTTGCAGATTTTTCACAACTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084  
Qy 1308 AGATCTGCTTGCAGATTTTTCACAACTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1367  
Db 1085 TGTCTTAAG 1144  
Qy 1368 TGTCTTAAG 1427  
Db 1145 ATGACTCCCACTAGTACTGCTGAG 1204  
Qy 1428 ATGACTCCCACTAGTACTGCTGAG 1487  
Db 1205 AACAGCGCAATGAGCTGGAAG 1264  
Qy 1488 AACAGCGCAATGAGCTGGAAG 1547  
Db 1265 TGTCTCAAAATGCAATTCAG 1324  
Qy 1548 TGTCTCAAAATGCAATTCAG 1607  
Db 1325 GCCCTTCAGTCCAG 1384



1088 CCGGACACACAGGCTGTCCAGATGTCTGCAACCGCCGAGTGGCACAGGCCCTC 1127  
Db 890 AGCAGTCTTTCGACAGAGTTCGCGCCACAGCAAGCTACGGGAGTCTTGTCTCTGC 949  
Qy 1128 CGGAGTCTTTCGACAGAGTTCGCGCCGAGCAAGCTACGGGAGTCTTGTCTCTGC 1187  
Db 990 CGGAGTCTTTCGACAGAGTTCGCGCGGAGCAAGCTACGGGAGTCTTGTCTCTCAA 1009  
Qy 1188 CGGAGTCTTTCGACAGAGTTCGCGGAGCAAGCTACGGGAGTCTTGTCTCTCAA 1247  
Db 1010 GAAG 1069  
Qy 1248 GAG 1307  
Db 1070 AGATCTGCGCTTCGAGATTTTTCACAGCTGCGAGCAGAGTCAAGTCTGTGACAG 1129  
Qy 1308 AGATCTGCGCTTCGAGATTTTTCACAGCTGCGAGCAGAGTCAAGTCTGTGACAG 1367  
1130 TGTCTTAAG 1189  
1368 TGTCTTAAG 1427  
Db 1190 ATGACTCCCACTACAGTACAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1249  
Qy 1428 ATGAG 1487  
Db 1250 AACAG 1309  
Qy 1488 AACAG 1547  
Db 1310 TGTCTTAAG 1369  
Qy 1548 TGTCTTAAG 1607  
Db 1370 GCGCTCCAGTACAG 1429  
Qy 1608 GCGCTCCAGTACAG 1667  
Db 1430 CCTCTGGGAG 1489  
Qy 1668 CCTCTGGGAG 1727  
Db 1490 AATTGAG 1549  
Qy 1728 AATTGAG 1787  
Db 1550 GATAGTATTCGAG 1609  
Qy 1788 AATGATATTCGAG 1844  
Db 1610 ATGCTGCTCTCTCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1669  
Qy 1845 ATGCTGCTCTCTCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1904  
Db 1670 GCTGCT 1729  
Qy 1905 TCCAGCT 1964  
Db 1730 GAAAG 1788  
Qy 1965 TGTAAAG 2023  
Db 1789 GTTTTAAAG 1844  
Qy 2024 GTTTTAAAG 2082  
Db 1845 TTAAG 1901  
Qy 2083 TTAAG 2142  
Db 1902 CCAAG 1961

Qy 2143 CCAAGTACAG 2202  
Db 1962 GCAAGCTGGGCTTATCATGTGTGTGATGATGATGATGATGATGATGATGATGAT 2015  
Qy 2203 GTAAG 2255

RESULT 7  
LOCUS AF014117 2549 bp mRNA ROD 20-NOV-1997  
DEFINITION Mus musculus GDNF receptor alpha (GDNFR-alpha) mRNA, complete cds.  
ACCESSION AF014117  
NID 92624960  
KEYWORDS house mouse.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 2549)  
AUTHORS Dey,B.K., Wong,Y.W. and Too,H.P.  
TITLE Cloning of a novel murine isoform of the glial cell line-derived neurotrophic factor receptor  
JOURNAL Neuroreport 9 (1) (1998) In press  
REFERENCE 2 (bases 1 to 2549)  
AUTHORS Dey,B.K., Wong,Y.W. and Too,H.P.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-1997) Biochemistry, 10 Kent Ridge Crescent, Singapore 119260, Singapore  
FEATURES  
source  
1. 2549  
location/Qualifiers  
/organism="Mus musculus"  
/strain="C57"  
/db\_xref="taxon:10090"  
/issue\_type="liver"  
1. 2549  
/gene="GDNFR-alpha"  
914..2320  
/gene="GDNFR-alpha"  
/note="glial cell line derived neurotrophic growth factor receptor"  
/codon\_start=1  
/product="GDNF receptor alpha"  
/db\_xref="PID:92624961"  
/translation="MFLATLYFVLPDLILMSAEVSGDRDLCVKASDGLKROSCST KYRTLQCVAGKRETNLSITSGLEAKDECSAMALOKSLYNCRKGRKKENCLRI YNSMYSLOGNDLLEDSPYEPVNSRSLDFRAVPIPSDPOQVEHISKNCNDAKA CNLDDCKIRSAVITPCTYSMSNEVCNKRCKAKAROFFDKVPAHSTGMLCSCRD VACTERRRQIVPVCYEEBERENCLNDDSKTNYICSRRLADFTNCOPEBSYSN CLKENYADCLAYSGLIGVMPNYSIDSSLSVAPWDCSNSNDLEDCLFLNFKD NMLKNALIOFGNSDVTMOPAPVQTTATTTAFRIKNRPLGPAENSEIEPTVL PRCANLOAKLKNVSGSTHLCISNDYDKDGLAGASHITTSMAAPPSCGLSLPV MFTLAALSLVSLAETS"

BASE COUNT 623 a 680 c 693 g 553 t  
ORIGIN  
Query Match 48.4% Score 1223; DB 23; Length 2549;  
Best Local Similarity 85.1% Pred. No. 0.00e+00;  
Matches 1608; Conservative 0; Mismatches 260; Indels 21; Gaps 13;

Db 665 CAGGTGGGTGGAAGTGAAGCCCTAAAGCGGGTCCGCTCCGCGCCCTCGCGCCGCTC 724  
Qy 291 CAGGTGGGTGGAAGTGAAGCCCTAAAGCGGGTCCGCTCCGCGCCCTCGCGCCGCTC 348  
Db 725 GGAGCTGAGTCCCTGG 784  
Qy 349 GGAGCTGAGTCCCTGG 407  
Db 785 ATGAGCTTAACTTGAAGTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 844  
Qy 408 ATGAGCTGAAGTCTGG 467  
Db 845 GCTCTCTCCGAGAGATCCGGTGGGCTTGGATTTT--GGGGGGGGGGGAGACAGCTG 901  
Qy 468 GCTCTCTCCGAGAGATCCGGTGGGCTTGGATTTTGGGGGGGGGGGAGACAGCTG 527

D 902 CCGGGGGGACATGTTCTAGCCACTGTACTTGTGCTGCCACTCTCTGATTTGCTG 961  
OY 528 CCGGGGGGACATGTTCTAGCCACTGTACTTGTGCTGCCACTCTCTGATTTGCTG 587  
D 962 ATGTGCGCCAGAGTGGTGGGGGACCGCTGGACTGTGTGAAGGACGATGATGATG 1021  
OY 588 CTGTGCGCCAGAGTGGTGGGGGACCGCTGGACTGTGTGAAGGACGATGATGATG 647  
D 1022 CTGAAGGACAGAGCTGACACACCAAGTACCGCAGCTGAGGAGCTGTGGGGGAG 1081  
OY 648 CTGAAGGACAGAGCTGACACACCAAGTACCGCAGCTGAGGAGCTGTGGGGGAG 707  
D 1082 GAAGCAACTTCAAGCTGACATCCGGCTGAGAGCCAGAGATGATGTCGCGAGGCTATG 1141  
OY 708 GAGCAACTTCAAGCTGACATCCGGCTGAGAGCCAGAGATGATGTCGCGAGGCTATG 767  
D 1142 GAGGCTTGAAGCAGAGTCTCTTACAGTGGCGCTGCAAGCGGGGCTAGAAAGAG 1201  
OY 768 GAGGCTTGAAGCAGAGTGGCTCTTACAGTGGCGCTGCAAGCGGGGCTAGAAAGAG 827  
D 1202 AAGATTTGTGCTATCTAAGAGATGATGAGCTGACAGAGCTGAGGAAATGACTACTG 1261  
OY 828 AAGATTTGTGCTATCTAAGAGATGATGAGCTGACAGAGCTGAGGAAATGACTACTG 887  
D 1262 GAAGATTTGTGCTATCTAAGAGATGATGAGCTGACAGAGCTGAGGAAATGACTACTG 1321  
OY 888 GAGATTTGTGCTATCTAAGAGATGATGAGCTGACAGAGCTGAGGAAATGACTACTG 947  
D 1322 TTGATATCAGATGTTTTCAGAGAGTGAACACATTTTCAAGAGGAAACATGCTGAT 1381  
OY 948 TTGATATCAGATGTTTTCAGAGAGTGAACACATTTTCAAGAGGAAACATGCTGAT 1007  
D 1382 GCAGCAAGGCTGCAACCTGATGACACTGCAAGAGTCCGCTACATCACC 1441  
OY 1008 GCAGCAAGGCTGCAACCTGATGACACTGCAAGAGTCCGCTACATCACC 1067  
D 1442 CCGGTGACCAAGAGTGTCAATGATGATGAGTGAAGCGGCAAGTGCACAAAGCCCTG 1501  
OY 1068 CCGGTGACCAAGAGTGTCAATGATGATGAGTGAAGCGGCAAGTGCACAAAGCCCTG 1127  
D 1502 AGGCAAGTTCCTGCAAGAGTTCAGAGGCAAGCAGAGTACGGGATGCTCTGCTGCTG 1561  
OY 1128 CCGGAGTTCTTTGAGAGTCCGCGGCAAGCAGAGTACGGGATGCTCTGCTGCTG 1187  
D 1562 CCGGAGTTCCTGCAAGAGTTCAGAGGCGGCAAGAGTATGCTCCCTGCTCTATGAA 1621  
OY 1188 CCGGAGTTCCTGCAAGAGTTCAGAGGCGGCAAGAGTATGCTCCCTGCTCTATGAA 1247  
D 1622 GAAGGAGAGGCGCAAGTCTGAATCTGCAAGAGTCTGCAAGAGTCTGCAAGAGTCTG 1681  
OY 1248 GAGAGGAGAGGCGCAAGTCTGAATCTGCAAGAGTCTGCAAGAGTCTGCAAGAGTCTG 1307  
D 1682 AGATCTGCGCTTGAAGATTTTTCACCAAGTTCAGAGGCAAGAGTCTGTCAGAAC 1741  
OY 1308 AGATCTGCGCTTGAAGATTTTTCACCAAGTTCAGAGGCAAGAGTCTGTCAGAAC 1367  
D 1742 TGTCTTAAGAGAGTACGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1801  
OY 1368 TGTCTTAAGAGAGTACGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1247  
D 1802 ATGATCTTAAGAGAGTACGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861  
OY 1428 ATGATCTTAAGAGAGTACGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1487  
D 1862 AACAGTGGCAATGAGCTGAGAGATGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 1921  
OY 1488 AACAGTGGCAATGAGCTGAGAGATGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 1547  
D 1922 TGTCTTAAGAGAGTACGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1981  
OY 1548 TGTCTTAAGAGAGTACGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1607

D 1982 GCCCCCGAGTCAGAGCAACCACTGCGAGACTACACCTGCTTCGGATCAAGAACAG 2041  
OY 1608 GCCCTCCAGTACAGACCACTGCGAGCTACACACCTGCGGGTAAAGAACAG 1667  
D 2042 CCTCTAGGCGAGAGCTGCTGAGTATGATTTCCACACAGCTTTTACACAGCTGCT 2101  
OY 1668 CCTCTAGGCGAGAGCTGAGTATGATTTCCACACAGCTTTTACACAGCTGCTG 1727  
D 2102 AATTGCGAGCAGAGAGTGAATCAATGATGCGGAGTACACATCTGTCTTCT 2161  
OY 1728 AATTGCGAGCAGAGAGTGAATCAATGATGCGGAGTACACATCTGTATTTCC 1787  
D 2162 GATTAATGAGTGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2221  
OY 1788 AATGTAATGAGTGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1844  
D 2222 ATGCTGCTCTCTCCAGCTGCGGCTGAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCT 2281  
OY 1845 ATGCTGCTCTCTCCAGCTGCGGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1904  
D 2282 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338  
OY 1905 TCCACCTTATCTTATCTTAAACAGAAACATCATGCTGCTTAAATAATCAATATGAGACA 1964  
D 2339 -GACAAAGA---AAGCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2394  
OY 1965 TGTAAAGAGCAAAACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2024  
D 2395 TTTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2450  
OY 2025 TTTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2083  
D 2451 TAAGAAAGCTTTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2510  
OY 2084 TAAGAAAGCTTTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2143  
D 2511 CAACACTCAGAAACCTGCGGAGTGTGT 2539  
OY 2144 CAACACTCAGAAAGCTTTGGGATATGCTGT 2172  
  
RESULT 8  
LOCUS AF015172 1651 bp mRNA ROD 20-NOV-1997  
DEFINITION Mus musculus GDNF receptor beta (GDNFR-beta) mRNA, complete cds.  
ACCESSION AF015172  
NID 92624962  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE  
1 (bases 1 to 1651)  
Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Dey,B.R., Wong,Y.W. and Too,H.P.  
TITLE Cloning of a novel murine isoform of the glial cell line-derived  
neurotrophic factor receptor  
JOURNAL Neuroreport 9 (1) (1998) In press  
2 (bases 1 to 1651)  
Dey,B.R., Wong,Y.W. and Too,H.P.  
AUTHORS Direct Submission  
TITLE Submitted (17-JUN-1997) Biochemistry, National University of  
Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore  
JOURNAL  
FEATURES  
source  
1..1651  
/organism="Mus musculus"  
/strain="C57"  
/db\_xref="taxon:10090"  
1..1651  
/gene="GDNFR-beta"  
230..1621  
/gene="GDNFR-beta"  
/note="glial cell-line derived neurotrophic factor  
receptor beta."  
/codon\_start=1







Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Euteria;  
Rodentia; Sciurgnathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 1392)  
Zhong,J.  
Direct Submission  
Submitted (16-OCT-1997) Zhong J., Molecular Neurobiochemistry  
NCT/172, Ruhr University Bochum, Universitaetsstr. 150, D-44780  
Bochum, GERMANY  
2 (bases 1 to 1392)  
Zhong,J., Amies,M. and Heumann,R.  
Rattus norvegicus GDNF-alpha/Tnfr1-delta mRNA, complete cds  
Unpublished  
Location/Qualifiers  
1. 1392  
/organism="Rattus norvegicus"  
/strain="Hanover"  
/db\_xref="taxon:10116"  
1. 1392  
/codon\_start=1  
/product="GDNF-alpha/Tnfr1-delta protein"  
/db\_xref="PID:e1154274"  
/db\_xref="PID:g2564302"  
/translation="MFLATLYFALPLDLMSAEVSGDRLDCVKASDQLEKQSCST  
KYRTLRQCVAGKEFESLTSLGLEAKDECRSMALKOKSLYCNCKRGMKKNCLRI  
YMSYOSLOGNDLLEDSPEYEVNRLSDIFRAVPEISVEHISKNNCLDAKACNLDD  
TCKYRSAYITPCTSMSEYCNKRCKHKLROFEDVPAKHSYMLFCSGRDIACIE  
RRROTIVPCYERERERPNCLSDSCNTYICSRILADPTNQPSPRSYSNLTEN  
YADCLAVSGLIGVTMPNTYDSSLSVAIPCDCNSGNDLEDLKLNFKDTCKL  
NAIDAFNGSDVTMOPAPVQTTATTATTAFRVKNRPLGPAAGSENIPIHVLPCAN  
LOAKLSNYSVSGTHLCLSDSEFGKLAGASSHITTKSMAPSPCSLSLPIVLMTA  
LAALLSYLAETS"

BASE COUNT 338 a 406 c 356 g 292 t  
ORIGIN

Query Match 38.4%; Score 970; DB 23; Length 1392;  
Best Local Similarity 88.2%; Pred. No. 0.00e+00;  
Matches 1223; Conservative 0; Mismatches 145; Indels 18; Gaps 6;

Db 1 ATGTCCTAGCCACTCTGACTTCGCGCTGCCACTCCCTGGATTCTGATGTCGCCCGAG 60  
Qy 540 ATGTCCTAGCCACTCTGACTTCGCGCTGCCACTCCCTGGATTCTGATGTCGCCCGAA 599  
Db 61 GTGAGTGGGAGACCGTGTGCTGTGTAAGGACGAGATCAGTGCCTGAAGGACAG 120  
Qy 600 GTGAGCGGGAGACCGCTGGATTGCGTGAAGACCAAGTCAATGCTGAGAGACAG 659  
Db 121 AGCTGACGACCAAGTACCGACACTAAGGACAGTGTGGCGGGGCAAGGAAACCACTTC 180  
Qy 660 AGCTGACGACCAAGTACCGACACTAAGGACAGTGTGGCGGGGCAAGGAAACCACTTC 719  
Db 181 AGCTGACATCGCGCTTGAGGCCAAGATAGTCCGTAGCGCATGAGGCCCTTGAAG 240  
Qy 720 AGCTGCGCATCGCGCTGAGGCCAAGGATGAGTCCGAGCGCATGAGGCCCTTGAAG 779  
Db 241 CAGAGTCTGTGTACAGCGCGCTGCAAGGGGAGTAAGAAAGAAAGAAATGTCTG 300  
Qy 780 CAGAGTCTGTGTACAGCGCGCTGCAAGGGGATGAGAAAGAAAGAAAGTGCCTG 839  
Db 301 CGTATCTACTGGAGCATGTACAGAGCCGTGCAAGGAAATGACTCTGTGAAAGATTCCCG 360  
Qy 840 CGCATTTACTGGAGCATGTACAGAGCCGTGCAAGGAAATGACTCTGTGAAAGATTCCCGA 899  
Db 361 TATGACCGCGTTACAGCAGGTTGTCAATATATTCGGGGCAGTCCGTTCAATACAG-T 419  
Qy 900 TATGACCGCGTTACAGCAGGATTGTCAATATATTCGGGGTGTCCCATTCATACAGAT 959  
Db 420 G-----G--A--CACATTCGAAAGGAAACATGCTGTGAGACACAGGCC 465  
Qy 960 GTTTTTCAGCAAGTGGAGCATTTCCCAAGGGAACATGCTGTGAGACACAGGCC 1019  
Db 466 TGCACACTGAGACACCTGTAAAGATGACAGTGGCGCTACATCACCCCTGCACACAC 525  
Qy 1020 TGCACACTGAGACATTTGCAAGAAAGTACAGTGGCGGTACATCACCCCGTGCACACAC 1079

Db 526 AGCATGTCGAAGAGGTGTGCAACCGCGCTGAAGGCCACAGAGGCCCTCAGCATTCTTC 585  
Qy 1080 AGCGTGTCAACGATGTCTGTGCAACCGCGCGGACAGAGTCCACAAAGGCCCTCCGAGATTCTT 1139  
Db 586 GACAAGTCTCCGGCCCAAGCAGACAGTACGAGGATGCTCTTGTCTCTCCGCGGACATCGCC 645  
Qy 1140 GACAAGTCTCCGGCCCAAGCAGACAGTACGAGGATGCTCTTGTCTCTCCGCGGACATCGCC 1199  
Db 646 TGCACGAGCGCGCGCGCAGACATATGTCTCCGCTGTCTCTATGAAGACGAGAGAG 705  
Qy 1200 TGCACAGAGCGGAGGACAGACAGTGTCTGTCTCTCTCTATGAAGAGGAGAGAG 1259  
Db 706 CCCAAGTCTGAGTGTGCAAGTCTGTGCAAGTCTGTGCAAGTCTGTGCAAGTCTGTGCAAG 765  
Qy 1280 CCCAAGTCTGAGTGTGCAAGTCTGTGCAAGTCTGTGCAAGTCTGTGCAAGTCTGTGCAAG 1319  
Db 766 GCAGATTTTATTTTACCAACTGCGCAGCAGAGTCAAGTCTGTGCAAGTCTGTGCAAGTCT 825  
Qy 1320 GCGGATTTTATTTTACCAACTGCGCAGCAGAGTCAAGTCTGTGCAAGTCTGTGCAAGTCT 1379  
Db 826 AACTACGACAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885  
Qy 1380 AACTACGCTGACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439  
Db 886 TACGTAGACGTCAGACAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 945  
Qy 1440 TACATAGACGTCAGACAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1499  
Db 946 GACCTGGAAGTCTGCTGGAATTTCTGGAATTTTCTGGAATTTTCTGGAATTTTCTGGAAT 1005  
Qy 1500 GACCTGGAAGTCTGCTGGAATTTCTGGAATTTTCTGGAATTTTCTGGAATTTTCTGGAAT 1559  
Db 1006 GCAATTCAGCTTTGGCAATGCTCAGATGTGACATGTGACATGTGACATGTGACATGTGACAT 1065  
Qy 1560 GCAATTCAGCTTTGGCAATGCTCAGATGTGACATGTGACATGTGACATGTGACATGTGACAT 1619  
Db 1066 CAGACACACTGCGCAC 1125  
Qy 1620 CAGACACACTGCGCAC 1679  
Db 1126 GCAGGCTGTGAGAAATGAGATCCACACACACACACACACACACACACACACACACACAC 1185  
Qy 1680 GCAGGCTGTGAGAAATGAGATCCACACACACACACACACACACACACACACACACACAC 1739  
Db 1186 CAGAACTGAAATCCATGTGTGCGGTGACACACACCTGTCTTCTGATGATGATTC 1245  
Qy 1740 CAGAACTGAAATCCATGTGTGCGGTGACACACACCTGTCTTCTGATGATGATTC 1799  
Db 1246 GGAAGATGCTGCTGCGGTGCTGCGGTGCTGCGGTGCTGCGGTGCTGCGGTGCTGCGGT 1305  
Qy 1800 GGAAGATGCTGCTGCGGTGCTGCGGTGCTGCGGTGCTGCGGTGCTGCGGTGCTGCGGT 1856  
Db 1306 CCCAGCTGCACTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1365  
Qy 1857 CCAAGCTGTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1916  
Db 1366 TCTGTA 1371  
Qy 1917 TCTTTA 1922

RESULT 11  
LOCUS GGU90541 3037 bp mRNA VMT 25-JUN-1997  
DEFINITION Gallus gallus GDNF receptor alpha (GDNFR1phn) mRNA, complete cds.  
ACCESSION U90541  
NID 92213802  
KEYWORDS chicken.  
SOURCE  
ORGANISM Gallus gallus  
Eukaryotae; mitochondria; eukaryotes; Metazoa; Chordata;  
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;  
Phasianidae; Phasianinae; Gallus.



AF038416:202..311,AF038417:132..266,AF038418:106..287,  
AF038419:118..171,80..832)  
/gene="GFR1"  
/product="GPI-linked anchor protein"  
order(AF038410:1..385,AF038411:1..529,AF038412:1..497,  
AF038413:1..363,AF038414:1..270,AF038415:1..596,  
AF038416:1..436,AF038417:1..397,AF038418:1..432,  
AF038419:1..256,1..832)  
/gene="GFR1"  
join(AF038411:357..396,AF038412:76..369,AF038413:162..245,  
AF038414:131..145,AF038415:147..483,AF038416:202..311,  
AF038417:132..266,AF038418:106..287,AF038419:118..171,  
80..226)  
/note="GFR1"  
/note="part of functional ligand for RET receptor tyrosine  
kinase"  
/codon\_start=1  
/product="GPI-linked anchor protein"  
/db\_xref="pid:92921545"  
/translation="MFLLYFALPLDLILLASVSGGDRDCVAKSDCKECSST  
KYRLROCAKRETFESLASGLEADECRSMENALKOKSLNCRKRMKREKNLRI  
TWSQSLQGNLEDESPYEPNRSLSDFRVPTSDVEQVEHIFGNCLDAAK  
CNLDICRKYRSAYITPCTSVSNDVNRRCNKLQFEDKVPKHSYGM,FCSRD  
IACERRRQITVPVCSYEEREPNCLNODCKTNYICRSRLADFTNCPESRSVSS  
CLKENYADCLAYSLIGVTMPNYIDSSLSVAPWCDCNSGNDLECLKFNFKD  
MTCNAIOAFNGNSDVTWQPAFPVOTTTATTTALRVKPKPLGPAENSEIPIHL  
PCANLQAKLKSNTSGNTHLCISNGNTEKELGASSHITTKSMAAPPSCGLSPILVL  
VVTALSTLSLETSS"  
80..832  
/gene="GFR1"  
/number=11  
227..832  
/gene="GFR1"  
BASE COUNT 245 a 160 c 167 g 260 t  
ORIGIN  
Query Match 19.4%; Score 491; DB 22; Length 832;  
Best Local Similarity 94.2%; Pred. No. 2,25e-297;  
Matches 733; Conservative 0; Mismatches 5; Indels 40; Gaps 2;  
exon  
3'UTR  
BASE COUNT 245 a 160 c 167 g 260 t  
ORIGIN  
Query Match 19.4%; Score 491; DB 22; Length 832;  
Best Local Similarity 94.2%; Pred. No. 2,25e-297;  
Matches 733; Conservative 0; Mismatches 5; Indels 40; Gaps 2;  
exon  
3'UTR  
BASE COUNT 245 a 160 c 167 g 260 t  
ORIGIN

AF012811 777 bp mRNA ROD 06-AUG-1997  
DEFINITION Mus musculus GDNF receptor alpha mRNA, partial cds.  
ACCESSION AF012811  
KEYWORDS 92286224  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;  
Murinae; Mus.  
REFERENCE 1 (bases 1 to 777)  
AUTHORS Pullitt,A. and Serl,M.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUL-1997) Laboratorio di Genetica Molecolare,  
Istituto Giannina Gaslini, Largo G. Gaslini, 5, Genova 16148, Italy  
FEATURES  
source  
1..777  
location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
1..430  
/organism="Mus musculus"  
/note="EST deposited under GenBank Accession Number W76716  
from Soares mouse embryo NM0E13.5 14.5 cDNA library; IMAGE  
Consortium"  
1..680  
/codon\_start=3  
/product="GDNF receptor alpha"  
/db\_xref="pid:92286225"  
/translation="LNQDSCKTNYICRSRLADFTNCPESRSVSNCLKENYADCL  
AYSLGIVMTPNYIDSSLSVAPWCDCNSGNDLECLKFNFKNTCLKNAIOF  
GNGSDVMTWQPAFPVOTTTATTTAFRIKKNPLDPASSEIPIHLVPCANLQAKL  
KSNVSGSTHLCLSDNDYCKDGLAGASSHITTKSMAAPPSCGLSLPVMVFTALALLS  
VSLAETS"  
BASE COUNT 202 a 212 c 174 g 189 t  
ORIGIN  
Query Match 19.3%; Score 488; DB 23; Length 777;  
Best Local Similarity 84.9%; Pred. No. 2,39e-295;  
Matches 661; Conservative 0; Mismatches 107; Indels 11; Gaps 7;  
Db 4 TGAATCTGCAGACCTCTGCAAGCAAAATTACATCTGCGATCTGCGATTTT 63  
Qy 1270 TGAATTTCCAGACCTCTGCAAGCAAAATTACATCTGCGATCTGCGATTTT 1329  
Db 64 TTACCACTGCCAGCAGCAAGTCTGTCAGCAACTGCTTAAGAGAACTACGAG 123  
Qy 1330 TTACCACTGCCAGCAGCAAGTCTGTCAGCAACTGCTTAAGAGAACTACGAG 1389  
Db 124 ACTGCTCTGCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183  
Qy 1390 ACTGCTCTGCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449

DB	184	CCAGCAGCCTCAGTGTGGGCGCGTGTGGATGTGGACGAAACAGTGCCATGACCTGGAA	243
Oy	1450	CCAGTAGCCTCAGTGTGGGCGCGTGTGGATGTGGACGAAACAGTGCCATGACCTGGAA	1509
Db	244	ATTGCTGAGTTTCTGAAATTTTTTAAAGGCAATACGTGTCTCAAAAAATGCAATTCAG	303
Oy	1510	AGTGCTGAAATTTTGAATTTCTTCAAGAGCAATACATGCTTTAAAAATGCAATTCAG	1569
Db	304	CCTTTGGCAATGGCTCGGATGTGACCATGTGGCAGCCAGCCCCCAGTCACGACACCA	363
Oy	1570	CCTTTGGCAATGGCTCGGATGTGACCATGTGGCAGCCAGCCCCCAGTCACGACACCA	1629
Db	364	CTGCCAGCACTACACTGCTCCGAGTCAAGCAAGCCTCTAGAGGCGAGAGGCTCTG	423
Oy	1630	CTGCCAGCACTACACTGCTCCGAGTCAAGCAAGCCTCTAGAGGCGAGAGGCTCTG	1689
Db	424	AGAAATGAGATTCACACACAGTTTAAACACCGCTGTCTAATTTGACGACAGAAAGCTGA	483
Oy	1690	AGAAATGAGATTCACACACAGTTTAAACACCGCTGTCTAATTTGACGACAGAAAGCTGA	1749
Db	484	AATCCAAATGATCGGGCAGTACACATCTGTCTTCTGATTAATGATTAACGAAAGATG	543
Oy	1750	AATCCAAATGATCGGGCAGTACACATCTGTCTTCTGATTAATGATTAACGAAAGATG	1809
Db	544	GTCCTGCTGTGCTCCACAGCCACATACACAAATCAATGCTGCTCTCCACAGCTGG	603
Oy	1810	GTCCTGCTGTGCTCCACAGCCACATACACAAATCAATGCTGCTCTCCACAGCTGG	1866
Db	604	GTCGAGCTCACTGCGCGGTGATGTGTTACCGCTCTGCTGCTGCTCTGTTGTTATCAT	663
Oy	1867	GTCGAGCTCACTGCGCGGTGATGTGTTACCGCTCTGCTGCTGCTCTGTTGTTATCAT	1926
Db	664	TGGCAGAA-ACAT-CGT-AGCTGCATCCAGTGTGAAA-GACAAAAGA---AAACCAAGT	716
Oy	1927	AAACATCATAGCTGATTAATAAAATACATATGACATGTAAAGACAAATACCAAGT	1986
Db	717	ATTCTGCTCCCTGCTCTCTTGTATATCTGAAATCCAGTTTAAAGCTGCTGTAGAA	775
Oy	1987	TATCTGTTCTGTTCTCTCTGTATATGCTGAATTCAGATTT-AGAGGCTCAGTTGGAA	2044

BASE COUNT	152 a	156 c	163 g	125 t
Query Match	13.48;	Score 339;	DB 22;	Length 596;
Best Local Similarity 100.0%;		Pred. No. 3,42e-195;		Mismatches 0;
Matches 339; Conservative		0;	Indels 0;	Gaps 0;
Db	145	AGTGAGGACGACATTTCCCAAGGAGGACCACTGCTGGATGACAGGAAGGCGCTGCAACTCGA	204	
Qy	971	AGTGAGGACGACATTTCCCAAGGAGGACCACTGCTGGATGACAGGAAGGCGCTGCAACTCGA	1030	
Db	205	CGACATTTTGCAAGAAGTACAGGTGGCGGTACATCACCCTGGACACCACGAGGTGTCCAA	264	
Qy	1031	CGACATTTTGCAAGAAGTACAGGTGGCGGTACATCACCCTGGACACCACGAGGTGTCCAA	1090	
Db	285	CGATGTCGCAACCGCGGCAAGTGGCACAAGGCCCTCGGAGATCTTTGACAAAGTCTCC	324	
Qy	1091	CGATGTCGCAACCGCGGCAAGTGGCACAAGGCCCTCGGAGATCTTTGACAAAGTCTCC	1150	
Db	325	GGCCAGGACGACGCTACGGAATGCTTTCTGCTGCTGCGGGACATCGCTGCACAGACG	384	
Qy	1151	GGCCAGGACGACGCTACGGAATGCTTTCTGCTGCTGCGGGACATCGCTGCACAGACG	1210	
Db	385	GAGGCGACAGACCATCGTGCCTGTGTCTCTATGAAAGAGGAGGAGGCCCACTGTTT	444	
Qy	1211	GAGGCGACAGACCATCGTGCCTGTGTCTCTATGAAAGAGGAGGAGGCCCACTGTTT	1270	
Db	445	GAATTTGCAGAGCTCTCTCAAGACGCAATTTACATCTGCAG	483	
Qy	1271	GAATTTGCAGAGCTCTCTCAAGACGCAATTTACATCTGCAG	1309	
RESULT	15			
LOCUS	HS6FRA1G01	385 bp	DNA	PRI
DEFINITION	Homo sapiens GPI-linked anchor protein (GFRAL) gene, exon 1.			10-APR-1998
ACCESSION	AF038410			
NID	g2921533			
KEYWORDS	1 of 11			
SEGMENT	human.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			
REFERENCE	1 (bases 1 to 385)			
AUTHORS	Angrist,M., Jing,S., Bolk,S., Bentley,K., Nallsamy,S.,			
TITLE	Human GFRAL: Cloning, mapping, genomic structure, and evaluation as			
JOURNAL	a candidate gene for Hirschsprung disease susceptibility			
REFERENCE	Genomics 48, 354-362 (1998)			
AUTHORS	2 (bases 1 to 385)			
TITLE	Angrist,M., Jing,S., Bolk,S., Bentley,K., Nallsamy,S.,			
JOURNAL	Halushka,M., Fox,G.M. and Chakravarti,A.			
FEATURES	Direct Submission			
SOURCE	Submitted (12-DEC-1997) Genetics, Case Western Reserve, 2109			
	Adelbert Road, Cleveland, OH 44106-4955, USA			
	Location/Qualifiers			
	1. 385			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="10"			
	/map="10q25"			
	1. 314			
	/gene="GFRAL"			
	/number=1			
BASE COUNT	83 a	111 c	91 g	100 t
ORIGIN				
Query Match	11.8%;	Score 298;	DB 22;	Length 385;
Best Local Similarity 100.0%;		Pred. No. 7.33e-168;		
Matches 298; Conservative		0;	Mismatches 0;	Indels 0;
			Gaps 0;	
Db	18	AATCTGCGCTCGAAGACGCGCATTTCTCCGGCGCGCTTCGATTAACCACTAATACCTCTAA	77	



Oy	1	AATCTGGCCTCGGAACAGCCATTCTCGGCGCCTTCGATACACCTAATACCTCTAA	60
Db	78	CGAGATCCGAGCCGAGGCGCTGCTCGGAATGTCCTGCGCCAACTGGGCTCTCGAG	137
Oy	61	CGAGATCCGAGCCGAGGCGCTGCTCGGAATGTCCTGCGCCAACTGGGCTCTCGAG	120
Db	138	CTCTCGAAGATTACCGCATCTATTTTTTTTTCTTTTTTCTTTTCTCAAGCGAGATAA	197
Oy	121	CTCTCGAAGATTACCGCATCTATTTTTTTTTCTTTTTTCTTTTCTCAAGCGAGATAA	180
Db	198	AGTGAGCCCGGAAAGGAGAGAGGGGCGGAGACACCATTTGCCCTGAAAGATTAATAAG	257
Oy	181	AGTGAGCCCGGAAAGGAGAGAGGGGCGGAGACACCATTTGCCCTGAAAGATTAATAAG	240
Db	258	TAAATAAACAACATGCGCTCTCGCGGACGCTGAGACGGGCTCGTTAGTCCAGATTGG	315
Oy	241	TAAATAAACAACATGCGCTCTCGCGGACGCTGAGACGGGCTCGTTAGTCCAGATTGG	298

```

arch completed: Wed Jun 24 19:01:51 1998
time : 5592 secs.

```

THIS PAGE BLANK (USPTO)

\*\*\*\*\*  
 W P S E I L I  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mparc\_h\_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Jun 24 20:29:12 1998; Maspar time 376.73 Seconds  
 868.165 Million cell updates/sec

ular output not generated.

Title: >US-08-866-354-1  
 Description: (1-2568) from US08866354.seq  
 Perfect Score: 2525  
 N.A. Sequence: 1 AATCGGCGCTGGGAACAGC.....CACCTCTGTACAAAGAAA 2568  
 Comp: TTAGACCGGAGACCTTGTCGCG.....GTGAGAACATCTTTCTTTT

Scoring table: TABLE default  
 Gap 6

Nmatch SMD : Dbase 0; Query 0

Searched: 176886 segs; 63680241 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: n-geneeq31-2

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
 14:part14 15:part15 16:part16 17:part17 18:part18  
 19:part19 20:part20 21:part21 22:part22 23:part23  
 24:part24 25:part25 26:part26 27:part27 28:part28  
 29:part29 30:part30 31:part31 32:part32 33:part33  
 34:part34 35:part35 36:part36 37:part37

Statistics: Mean 9.999; Variance 6.681; scale 1.497

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1137	45.0	2378	37	T84975	Rat glial cell derive
2	612	24.2	840	37	T84977	Mouse GDNFR alpha clo
3	226	9.0	418	37	T84976	Mouse GDNFR alpha clo
4	131	5.2	453	37	T84979	Human EST-derived seq
5	124	4.9	351	37	T84978	Human EST-derived seq
6	96	3.8	385	15	O85526	DNA probe 32 detects
7	90	3.6	201	37	T84981	Human EST-derived seq
8	90	3.6	201	37	T84980	Human EST-derived seq
9	45	1.8	204	1	N81164	Base substituted E.co
10	44	1.7	91	9	O51746	Oligonucleotide probe
11	42	1.7	91	9	O51746	Oligonucleotide probe
12	43	1.7	204	1	N81164	Base substituted E.co
13	39	1.5	114	12	O70469	Generic DNA sequence
14	38	1.5	114	12	O70467	Generic DNA sequence

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
15	39	1.5	172	32	T76363	Human Interleukin 8 a
16	36	1.4	114	12	O70468	Generic DNA sequence
17	36	1.4	114	12	O70465	Generic DNA sequence
18	36	1.4	114	12	O70466	Generic DNA sequence
19	36	1.4	114	12	O70467	Generic DNA sequence
20	36	1.4	114	12	O70468	Generic DNA sequence
21	35	1.4	114	12	O70465	Generic DNA sequence
22	35	1.4	114	12	O70469	Generic DNA sequence
23	34	1.3	114	12	O70470	Generic DNA sequence
24	34	1.3	114	12	O70471	Generic DNA sequence
25	33	1.3	114	12	O70472	Generic DNA sequence
26	32	1.3	114	12	O70473	Generic DNA sequence
27	32	1.3	114	12	O70474	Generic DNA sequence
28	33	1.3	114	12	O70470	Generic DNA sequence
29	32	1.3	168	31	T76270	Human MDNCF antisense
30	32	1.3	172	32	T76363	Human Interleukin 8 a
31	32	1.3	178	31	T76405	Human endocytelin-1 an
32	32	1.3	190	31	T76452	Chymase antisense o
33	30	1.2	114	12	O70472	Generic DNA sequence
34	30	1.2	178	31	T76405	Human endocytelin-1 an
35	30	1.2	250	19	T22691	Human gene signature
36	30	1.2	8920	11	O62924	Carbamoyl-phosphate
37	29	1.1	114	12	O70473	Generic DNA sequence
38	27	1.1	128	31	T76233	Human IL6 antisense o
39	27	1.1	168	31	T76270	Human MDNCF antisense
40	28	1.1	416	8	O59427	Human brain expressed
41	29	1.1	598	1	N80919	Type A insertion gene
42	29	1.1	1021	16	T04613	5' flanking region of
43	29	1.1	2454	15	T01567	Hamster cDNA probe sp
44	27	1.1	2504	30	T62359	Schizosaccharomyces p
45	29	1.1	4641	1	N81538	Sequence of a gene fr

#### ALIGNMENTS

RESULT 1  
 ID T84975 standard; cDNA; 2378 BP.  
 AC T84975;  
 DE 27-Apr-1998 (first entry)  
 DE Rat glial cell derived neurotrophic factor receptor alpha cDNA.  
 KW Glial cell derived neurotrophic factor receptor alpha; GDNFR alpha;  
 OS GDNFR; rat; Kidney disease; glomerulonephritis; therapy; ds.  
 OS Rattus sp.  
 FH Key Location/Qualifiers  
 FT CDS 117..1523  
 FT sig\_peptide 117..188  
 FT mat\_peptide 189..1520  
 FT FT tag- b  
 FT mat\_peptide 189..1520  
 FT FT tag- c  
 PN W09733912-A2.  
 PD 18-SEP-1997.  
 PF 13-MAR-1997; U04363.  
 PR 14-MAR-1996; US-618236.  
 PR 14-MAR-1996; US-615902.  
 PA (GETH ) GENENTECH INC.  
 PI Klein RD, Moore MW, Rosenthal A, Ryan AM;  
 DR WPI; 97-470819/43.  
 DR P-PSDB; W27327.  
 PT Isolated glial cell derived neurotrophic factor receptor alpha -  
 PT useful to develop products to diagnose and treat associated  
 PT disorders, particularly enteric nervous system or kidney disorders  
 PS C14m 23; Page 74-77; 100pp; English.  
 CC This cDNA codes for full-length rat glial cell derived neurotrophic  
 CC factor receptor alpha (GDNFR alpha) (see W27327), a novel  
 CC GPT-linked protein that is a ligand-binding component of the  
 CC receptor system for GDNF. It was isolated by expression cloning.  
 CC A cDNA library obtained from ventral midbrain tissue of E14 rat  
 CC embryos was generated in a cytomegalovirus-based vector. cDNA  
 CC clones were transfected into COS 7 cells and expression of  
 CC putative GDNF receptors was detected by binding of iodinated GDNF.  
 CC A single positive pool was obtained, from which the cDNA clone was  
 CC isolated. An expression vector containing the cDNA can be used to



DE Mouse GDNFR alpha clone 26 3' end.  
 KW G1a1 cell derived neurotrophic factor receptor alpha: GDNFR alpha;  
 KW GDNFR: mouse: kidney disease; glomerulonephritis; therapy; ss.  
 OS Mus musculus.  
 PN W09733912-A2.  
 PD 18-SEP-1997.  
 PF 13-MAR-1997; U04363.  
 PR 14-MAR-1996; US-618236.  
 PR 14-MAR-1996; US-615902.  
 PA (GETH ) GENENTECH INC.  
 PI Klein RD, Moore MM, Rosenthal A, Ryan AM;  
 DR WPI: 97-470819/43.  
 PT Isolated glial cell derived neurotrophic factor receptor alpha -  
 PT useful to develop products to diagnose and treat associated  
 PT disorders, particularly enteric nervous system or kidney disorders  
 PS Example 1, Page 60-61, 100pp; English.  
 CC This cDNA sequence comprises the 3' end of mouse full-length  
 CC glial cell derived neurotrophic factor receptor alpha (GDNFR alpha)  
 CC (see W27327) clone 26; the 5' end of the clone is given in  
 CC T84978. Clone 26 was isolated from a mouse cDNA library using  
 CC rat GDNFR alpha cDNA (see T84975) as probe. The invention relates  
 CC to novel uses of GDNFR and its receptor. In particular, it relates  
 CC to native rat GDNFR alpha (see W27327), its variants and soluble  
 CC derivatives (extracellular domain), chimeric GDNFR alpha and  
 CC antibodies which bind to the GDNFR alpha, including agonist and  
 CC neutralising antibodies, as well as various uses for these  
 CC molecules. It also relates to assay systems for detecting ligands  
 CC to GDNFR alpha, systems for studying the physiological role of  
 CC GDNFR, diagnostic techniques for identifying GDNFR-related conditions,  
 CC methods for identifying molecules homologous to GDNFR alpha, and  
 CC therapeutic techniques (claimed) for the treatment of GDNFR-related  
 CC and GDNFR alpha-related conditions, particularly kidney disease  
 CC associated with glomerulonephritis and enteric nervous system  
 CC related disorders. Transgenic and knockout animals are also  
 CC claimed.  
 CC Sequence 840 BP; 205 A; 245 C; 200 G; 190 T;  
 SO

Query Match 24.2%; Score 612; DB 37; Length 840;  
 Best Local Similarity 88.0%; Pred. No. 0.00e+00;  
 Matches 726; Conservative 0; Mismatches 96; Indels 3; Gaps 1;

Db 1 cgcgcgaagtcgcacaaagccctcagcagctcttcgcacaaagtcgcacaaagcagc 60  
 Oy 1104 cgcgcgaagtcgcacaaagccctcagcagctcttcgcacaaagtcgcacaaagcagc 1163

Db 61 taagcgatgct 120  
 Oy 1164 taagcgatgct 1223

Db 121 atcgccccgtgtct 180  
 Oy 1224 atcgccccgtgtct 1283

Db 181 tccctcaagaacaattatcctcgcagatcgcctctgcagatcttttaccacagtcag 240  
 Oy 1284 tccctcaagaacaattatcctcgcagatcgcctctgcagatcttttaccacagtcag 1343

Db 241 ccagagtcgaagtcctctcgcagcaactcttaagaggaactacgcagcagctcctctgagc 300  
 Oy 1344 ccagagtcgaagtcctctcgcagcaactcttaagaggaactacgcagcagctcctctgagc 1403

Db 301 tectctggagatgattggcagcaatcactactactactactactactactactactactact 360  
 Oy 1404 tectctggagatgattggcagcaatcactactactactactactactactactactactact 1463

Db 361 gtggcgccgtgtgtgcagtcagcaacagtcgcaatgagtcgagtcgagtcgagtcgagtc 420  
 Oy 1464 gtggcgccgtgtgtgcagtcagcaacagtcgcaatgagtcgagtcgagtcgagtcgagtc 1523

Db 421 ctgaatttttttaaggacatacgtgtctcaaaaatgcaattcgaagcctttggcaattgac 480  
 Oy 1524 ttgaatttttttaaggacatacgtgtctcaaaaatgcaattcgaagcctttggcaattgac 1583

Db 481 tcgagatgacacatgtgacagccagcccccagctccagacacacacactgcccagactacc 540  
 Oy 1584 tcgagatgacacatgtgacagccagcccccagctccagacacacacactgcccagactacc 1643

Db 541 actgcttcggatcaagaacacagcctcctagggccagcagcagctctgagatgagatcc 600  
 Oy 1644 actgcttcggatcaagaacacagcctcctagggccagcagcagctctgagatgagatcc 1703

Db 601 aacacgcttttaccacagctgtgctcatttgaggcacagaaactgaaatcgaatgtatcg 660  
 Oy 1704 aacacgcttttaccacagctgtgctcatttgaggcacagaaactgaaatcgaatgtatcg 1763

Db 661 ggcagtaacacatctctgtctctctcgtatgatactacgaaagagtcgtctgtgtgccc 720  
 Oy 1764 ggcagtaacacatctctgtctctcgtatgatactacgaaagagtcgtctgtgtgccc 1820

Db 721 tccagccacataccacacaaataatcagtcgctcctccagctcgtgtgagctactg 780  
 Oy 1821 tccagccacataccacacaaataatcagtcgctcctccagctcgtgtgagctactg 1880

Db 781 ccggtgagtggtgttcacagctcgtgctgctgctgtgtgtgtgtgtgtgtgtgtgtgtgt 825  
 Oy 1881 ccggtgagtggtgtgttcacagctcgtgctgctgctgtgtgtgtgtgtgtgtgtgtgtgtgt 1925

RESULT 3  
 ID T84976 standard; cDNA: 418 BP.  
 AC T84976.  
 DT 27-APR-1998 (first entry)  
 DE Mouse GDNFR alpha clone 26 5' end.  
 KW G1a1 cell derived neurotrophic factor receptor alpha: GDNFR alpha;  
 KW GDNFR: mouse: kidney disease; glomerulonephritis; therapy; ss.  
 OS Mus musculus.  
 FH Key  
 FT CDS Location/Qualifiers  
 FT CDS 287..418  
 FT CDS /tag= "a  
 FT CDS /note= "5' end of GDNFR alpha coding region"  
 PN W09733912-A2.  
 PD 18-SEP-1997.  
 PF 13-MAR-1997; U04363.  
 PR 14-MAR-1996; US-618236.  
 PR 14-MAR-1996; US-615902.  
 PA (GETH ) GENENTECH INC.  
 PI Klein RD, Moore MM, Rosenthal A, Ryan AM;  
 DR WPI: 97-470819/43.  
 PT Isolated glial cell derived neurotrophic factor receptor alpha -  
 PT useful to develop products to diagnose and treat associated  
 PT disorders, particularly enteric nervous system or kidney disorders  
 PS Example 1, Page 60; 100pp; English.  
 CC This cDNA sequence comprises the 5' end of mouse full-length  
 CC glial cell derived neurotrophic factor receptor alpha (GDNFR alpha)  
 CC (see W27327) clone 26; the 3' end of the clone is given in  
 CC T84977. Clone 26 was isolated from a mouse cDNA library using  
 CC rat GDNFR alpha cDNA (see T84975) as probe. The invention relates  
 CC to novel uses of GDNFR and its receptor. In particular, it relates  
 CC to native rat GDNFR alpha (see W27327), its variants and soluble  
 CC derivatives (extracellular domain), chimeric GDNFR alpha and  
 CC antibodies which bind to the GDNFR alpha, including agonist and  
 CC neutralising antibodies, as well as various uses for these  
 CC molecules. It also relates to assay systems for detecting ligands  
 CC to GDNFR alpha, systems for studying the physiological role of  
 CC GDNFR, diagnostic techniques for identifying GDNFR-related conditions,  
 CC methods for identifying molecules homologous to GDNFR alpha, and  
 CC therapeutic techniques (claimed) for the treatment of GDNFR-related  
 CC and GDNFR alpha-related conditions, particularly kidney disease  
 CC associated with glomerulonephritis and enteric nervous system  
 CC related disorders. Transgenic and knockout animals are also  
 CC claimed.  
 CC Sequence 418 BP; 70 A; 119 C; 150 G; 79 T;  
 SO

Query Match 9.0%; Score 226; DB 37; Length 418;  
 Best Local Similarity 83.3%; Pred. No. 1.22e-117;  
 Matches 320; Conservative 0; Mismatches 58; Indels 6; Gaps 4;

Db 38 cagggtgggtcggagctgacccctaaagcgggtccgctccgctcgccgcgcgc 97  
|||  
Oy 291 CAGGTGGGTGGAGCTACCCCTAAAGCGAACCGCTCCGCTCCCTCCATC-C-C 348  
|||  
Db 98 ggaactgagtcctcggcggtggtggcgagagcaacgggagctctcaccctgg 157  
|||  
Oy 349 GGAGCTGATGGCGGGCGGGGTGG-CGTGCTGCCAGACCGGAGTTCTCTTTACACTGG 407  
|||  
Db 158 atgagagcttaactttagtggccagagagcgagtcgcccgggagtcctcagcgtga 217  
|||  
Oy 408 ATGAGCTGAATCTTGGCGGCGGAGCAGACAGCTGCTCGGGGATCGCTGACGCTGA 467  
|||  
Db 218 gctctctcggagagtcgggtggcgcttgattt---gggggggggggagacagctg 274  
|||  
Oy 468 GCTCCCTCGGAGACCGGAGCGGCGGCTCGGGATTTTGGGGGGGGGGGACCAACGCC 527  
|||  
Db 275 cgcggtgagcacaatgtctcagccactctgactctgctgctcactcctgattgctg 334  
|||  
Oy 528 CGCGCGGCGGACCATGTTCTTCTGGCGACCTGTACTTGGCGGCTTGGACTTGTCTC 587  
|||  
Db 335 atgtcgcgcgagtgatgtgtgggagccgctgagctgtgtgaagccagtatcagtcg 394  
|||  
Oy 588 CTGTGCGGCGGAGTGAGCGGCGGAGACCGCTGATTCGTGAAGCCAGATCATCATGTC 647  
|||  
Db 395 ctgaagagacagagctgagcacc 418  
|||  
Oy 648 CTGAAGAGCAGAGCTGCACACCC 671  
|||  
RESULT 4  
ID T84978 standard; DNA; 453 BP.  
AC T84979:  
DE 27-APR-1998 (first entry)  
KW Human EST-derived sequence y170a10.r1.  
KW Glial cell derived neurotrophic factor receptor alpha; GDNFR alpha;  
KW GDNF; human; kidney disease; glomerulonephritis; therapy;  
KW expressed sequence tag; EST; ss.  
OS Homo sapiens.  
PN MO9733912-A2.  
PD 18-SEP-1997.  
PE 13-MAR-1997; U04363.  
PR 14-MAR-1996; US-618236.  
PR 14-MAR-1996; US-615902.  
PA (GETH ) GENENTECH INC.  
PI Klein RD, Moore MM, Rosenthal A, Ryan AM;  
WP1: 97-470819/43.  
PT Isolated glial cell derived neurotrophic factor receptor alpha -  
PT useful to develop products to diagnose and treat associated  
PT disorders, particularly enteric nervous system or kidney disorders  
PS Example 1: Page 61: 100pp: English.  
CC This DNA sequence comprises a human EST-derived sequence designated  
CC y170a10.r1. It can be used to identify glial cell derived  
CC neurotrophic factor receptor (GDNFR) sequences, including human  
CC variants. Also suitable as probes for GDNFR sequences are another  
CC human EST-derived sequence designated ye83h05.r1 (see T84978), and  
CC fragments of these EST-derived sequences (see T84980-81) or  
CC proteins encoded by them. The invention relates to novel uses of  
CC GDNF and its receptor. In particular, it relates to native rat  
CC GDNFR alpha (see W27327), its variants and soluble derivatives  
CC (extracellular domain), chimeric GDNFR alpha and antibodies which  
CC bind to the GDNFR alpha, including agonist and neutralising  
CC antibodies, as well as various uses for these molecules. It also  
CC relates to assay systems for detecting ligands to GDNFR alpha,  
CC systems for studying the physiological role of GDNF, diagnostic  
CC techniques for identifying GDNF-related conditions, methods for  
CC identifying molecules homologous to GDNFR alpha, and claimed  
CC methods for the treatment of GDNF-related and GDNFR alpha-related  
CC conditions, particularly kidney disease associated with  
CC glomerulonephritis and enteric nervous system related disorders.  
CC Transgenic and knockout animals are also claimed.  
SQ Sequence 453 BP; 86 A; 152 C; 121 G; 91 T;

Query Match 5.2%; Score 131; DB 37; Length 453;  
Best Local Similarity 71.4%; Pred. No. 3,49e-59;  
Matches 242; Conservative 0; Mismatches 94; Indels 3; Gaps 3;  
Db 3 aaccattgcttgatgtctgccaagagccttgcaactgatatgcaactgccaagagctgctgc 62  
|||  
Oy 993 AACAACTGCTGGATGACGAGGAAGGCGCTGCACACTGCACACATTTTGCAGAGTACAG 1052  
|||  
Db 63 tctctcacaatcctcactcgtcaaacgcgagatctgcaccacgagcgctcgaacgcgcg 122  
|||  
Oy 1053 TCGGCGTACATCACCCCGTCCA-CCACACACG-T-GTCCACAGATGTCTCAACCGCGCG 1109  
|||  
Db 123 aagtcaccaagagccctgcgcccagttcttgaccgggtgcccagagtgaccatacgc 182  
|||  
Oy 1110 AAGTGCCACAAAGGCGCTCCGCACTTTTGACAAAGTCCCGGCCAAGCAGCTACGGA 1169  
|||  
Db 183 atgtctcttctctctcctcgaagaccagcgctgctgagtcgctgcggtgcaaacatcctg 242  
|||  
Oy 1170 ATGCTTTCGTCTCTCTCGCGGACATGCTGCACAGCGGAGGCGACAGACCATGCTG 1229  
|||  
Db 243 cccagctgctcctcatgaggaagagagcccaactgctgagctgctgctgctgctg 302  
|||  
Oy 1230 CTTGTGCTCTCTATGAAAGAGAGGACCAACCAACTGTTCATTTGACAGACTCTCTGC 1289  
|||  
Db 303 cgcagtcagcaccactgtgtgctcccggtcngcgaactt 341  
|||  
Oy 1290 AAGCAGATTTCATCTGCAGATCTCGCTTGGGATTTT 1328  
|||  
RESULT 5  
ID T84978 standard; DNA; 351 BP.  
AC T84978:  
DE 27-APR-1998 (first entry)  
KW Human EST-derived sequence ye83h05.r1.  
KW Glial cell derived neurotrophic factor receptor alpha; GDNFR alpha;  
KW GDNF; human; kidney disease; glomerulonephritis; therapy; EST;  
KW expressed sequence tag; ss.  
OS Homo sapiens.  
PN MO9733912-A2.  
PD 18-SEP-1997.  
PE 13-MAR-1997; U04363.  
PR 14-MAR-1996; US-618236.  
PR 14-MAR-1996; US-615902.  
PA (GETH ) GENENTECH INC.  
PI Klein RD, Moore MM, Rosenthal A, Ryan AM;  
WP1: 97-470819/43.  
PT Isolated glial cell derived neurotrophic factor receptor alpha -  
PT useful to develop products to diagnose and treat associated  
PT disorders, particularly enteric nervous system or kidney disorders  
PS Example 1: Page 61: 100pp: English.  
CC This DNA sequence comprises a human EST-derived sequence designated  
CC ye83h05.r1. It can be used to identify glial cell derived  
CC neurotrophic factor receptor (GDNFR) sequences, including human  
CC variants. Also suitable as probes for GDNFR sequences are another  
CC human EST-derived sequence designated y170a10.r1 (see T84978), and  
CC fragments of these EST-derived sequences (see T84980-81) or  
CC proteins encoded by them. The invention relates to novel uses of  
CC GDNF and its receptor. In particular, it relates to native rat  
CC GDNFR alpha (see W27327), its variants and soluble derivatives  
CC (extracellular domain), chimeric GDNFR alpha and antibodies which  
CC bind to the GDNFR alpha, including agonist and neutralising  
CC antibodies, as well as various uses for these molecules. It also  
CC relates to assay systems for detecting ligands to GDNFR alpha,  
CC systems for studying the physiological role of GDNF, diagnostic  
CC techniques for identifying GDNF-related conditions, methods for  
CC identifying molecules homologous to GDNFR alpha, and claimed  
CC methods for the treatment of GDNF-related and GDNFR alpha-related  
CC conditions, particularly kidney disease associated with  
CC glomerulonephritis and enteric nervous system related disorders.  
CC Transgenic and knockout animals are also claimed.  
SQ Sequence 351 BP; 77 A; 117 C; 96 G; 57 T;  
Query Match 4.9%; Score 124; DB 37; Length 351;



```

Best Local Similarity 72.7%: Pred. No. 5,78e-55; Matches 216; Conservative 0; Mismatches 78; Indels 3; Gaps 3;
Dy 40 ccaagagaacattcgtctgtagtgcgccaaagccctgaacctgatgaacaatcaga 99
Oy 985 CCAAGAGGAACACTGCTGGATGCAGGAAAGCGCTGCAACTCGACAGACTTGGCAGA 1044
Db 100 agctgcgtcctcttcatactccatctcgaaacgcggagatctcgccaacgcgtcga 159
Oy 1045 AGTACAGGTCCGGCGATCATCACCCCCGTCCA-CACACAGCG-T-GTCCACGATGTCGA 1101
Db 160 accgcgcgaagtgcacaaaggccctcgccagactcttcgcaccgggtgcccaagagaca 219
Oy 1102 ACCGCCGAAAGTGCCACAGAGCGCTTCGCGCAATTCTTTGACAAAGTCCCGCCAAAGCACA 1161
Db 220 cctaaccgatctctctctgcctctgcgaagaatcaaagcgtgcgtcgcagcnvcgagncaaa 279
Oy 1162 GCTAACGGAATGCTCTTTCGTCTTCGCGCGGACATTCGCTGCACAGAGGGAGCGGACAGAGA 1221
Db 280 ccattctgcaccaactgtctctcatatgaagaagaagagaaaccactgcnttgacntgc 336
Oy 1222 CCATCTGTCCTCCTGTGCTCTCATGAGAGAGGAGGAGAGCCCACTGTTGTAATTGTC 1278

RESULT 6
ID T84981 standard; DNA: 385 BP.
AC Q85526;
DE 10-JAN-1996 (first entry)
DB DNA probe 32 detects DNA-protein complex in immortal cells.
KW DNA-protein complex; detection; proliferation; tumour formation;
KM diagnose; malignancy; biopsy; probe; ss.
OS Synthetic.
PN WO9502701-A1.
PD 26-JAN-1995.
PE 13-JUL-1994; E02307.
PR 15-JUL-1993; DE-323727.
PI AbXen H, Albert W, Jungfer H, AbXen HJ;
PA (BOE ) BOEHRINGER MANNHEIM GMBH.
DR WP1: 95-067344/09.
PT New DNA-protein complex characteristic of cells with unlimited
proliferation capacity - and its components and derived antibodies,
PT useful in diagnosing malignant tumours.
PS Claim 1, Page 73; 106pp. German.
CC The DNA shown is found in human or animal cells that have an unlimited
capacity for unlimited cell proliferation or tumour formation. They have
CC no ability to promote immortalisation of the cells, and are usually
CC found in a DNA-protein complex in the cell cytoplasm. The DNA is useful
for detection of these complexes, and diagnosis of malignant tumours.
CC Differentiation between cells with unlimited and only transient
proliferation is possible. (See also Q85493-54)
Sequence 385 BP; 121 A; 83 C; 90 G; 91 T;

Query Match 3.8%; Score 96; DB 15; Length 385;
Best Local Similarity 92.9%; Pred. No. 2.63e-38;
Matches 104; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Dy 1 gatctgcctgcagattttttaccacactccagccagcagatcaaggtctcgcagcaact 60
Oy 1309 GATCGCGCTCGCGGATTTTTTACCAACTCCAGCCAGCCAGACTCAAGGTCTGTACAGCT 1368
Db 61 gttctaagagaactaagcagactgctctcgtgacctactcggagactattgg 112
Oy 1369 GTCTAAGGAATACTACCTGACTGCTCCCTCCGCTACTCGGGGCTATTGG 1420

RESULT 7
ID T84981 standard; DNA: 201 BP.
AC T84981;
DE 27-APR-1998 (first entry)
DB Human EST-derived sequence y170a10.r1 fragment.
KW Gli1 cell derived neurotrophic factor receptor alpha; GDNF; human; kidney disease; glomerulonephritis; therapy; EST;
expressed sequence tag; ss.
```

OS Homo sapiens.  
PN W09733912-A2.  
PD 18-SEP-1997.  
PF 13-MAR-1997; U04363.  
PR 14-MAR-1996; US-618236.  
PR 14-MAR-1996; US-615902.  
PA (GENET ) GENENTECH INC.  
PI Klein RD, Moore MW, Rosenthal A, Ryan AM;  
WP1: 97-470819/43.  
PT Isolated glial cell derived neurotrophic factor receptor alpha -  
PT useful to develop products to diagnose and treat associated  
PT disorders, particularly enteric nervous system or kidney disorders  
PS Example 1, Page 61; 100pp: English.  
CC This DNA sequence comprises a fragment of a human EST-derived  
CC sequence designated y10a10.r1 (see T84979). It can be used as  
CC a probe to identify glial cell derived neurotrophic factor receptor  
CC (GDNFR) sequences, including human variants. The invention relates  
CC to novel uses of GDNF and its receptor. In particular, it relates  
CC to rat GDNFR alpha (see W27337), its variants and soluble derivatives  
CC (extracellular domain), chimeric GDNFR alpha and antibodies which  
CC bind to the GDNFR alpha, including agonist and neutralising  
CC antiodies, as well as various uses for these molecules. It also  
CC relates to assay systems for detecting ligands to GDNFR alpha.  
CC systems for studying the physiological role of GDNF, diagnostic  
CC techniques for identifying GDNF-related conditions, methods for  
CC identifying molecules homologous to GDNFR alpha, and claimed  
CC methods for the treatment of GDNF-related and GDNFR alpha-related  
CC conditions, particularly kidney disease associated with  
CC glomerulonephritis and enteric nervous system related disorders.  
CC Transgenic and knockout animals are also claimed.  
SQ Sequence 201 Bp; 41 A; 78 C; 46 G; 36 T;

```

Query Match: 3.6%; Score 90; DB 37; Length 201;
Best Local Similarity 76.4%; Pred. No. 8,55e-35;
Matches 152; Conservative 0; Mismatches 44; Indels 3; Gaps 3;

Db      3  aacccttcctcggatgtcgtcccaaggccctgcacacctgatatcaactgcagaagtcgc 62
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      993 AACACATCCTCGTAGATGACGAGCAAGCGCTGCACACTCGACACATTTGCAGAGTACAG 1052

Db      63  tctctctacatctccatctgcacccgcgagatctctgccaccagagcgtcgaacgcgcg 122
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1053 TCGGGGTCAATCACCCCGCTGCA-CCACAGAG-T-GTCCAAAGATGTCTGCACACCGCGC 1109

Db      123  aagttcccaaggccctcgcgcagttcttcgcacgggttcgcccaagcagatcaactaccg 182
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1110 AAGTCCCAAGAGGCCCTTCGCGCAATTTTGACAAAGTCCGGCCAAAGCACAGCTACGGA 1169

Db      183  atgcctctctgcctcctgc 201
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1170 ATGCTCTCTGCTGCTCCTGCC 1188

RESULT      8
ID      T84980 standard; DNA; 201 BP.
AC      T84980;
DT      27-APR-1998 (first entry)
DE      Human EST-derived sequence y83h05.r1 fragment.
KW      Glnal cell derived neurotrophic factor receptor alpha; GDNFR alpha;
KW      GDNF; human; kidney disease; glomerulonephritis; therapy; EST;
KW      expressed sequence tag; ss.
OS      Homo sapiens.
PN      WO9733912-A2.
PD      18-SEP-1997.
PE      13-MAR-1997; U04363.
PR      14-MAR-1996; US-618236.
PR      14-MAR-1996; US-615902.
PA      (GETH ) GENENTECH INC.
PI      Klein RD, Moore MW, Rosenthal A, Ryan AM;
PI      Wpi: 97-470819/43.
PT      Isolated glnal cell derived neurotrophic factor receptor alpha -
PT      useful to develop products to diagnose and treat associated
PT      disorders, particularly enteric nervous system or kidney disorders

```



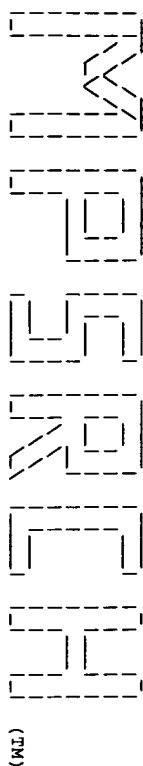
KW TSR: totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.

[illegible]

PT adenosine-free antisense oligo:nucleotide to airway epithellum of  
PT subject  
PS Claim 5; Page 36; 71pp; English.  
CC A method for treating airway disease in a subject has been produced,  
CC which involves the topical administration of an essentially adenosine  
CC free antisense oligonucleotide (ON) to the airway epithellum of the  
CC subject. The present sequence is an antisense oligonucleotide specific  
CC for the human interleukin 8, targeted at the initiation codon. The  
CC method can be used to treat airway diseases such as cystic fibrosis,  
CC asthma, chronic obstructive pulmonary disease, bronchitis and other  
CC airway diseases characterised by an inflammatory response. By  
CC eliminating adenosine from the antisense ON, its liberation upon  
CC antisense degradation is prevented, thereby preventing adenosine-  
CC induced bronchoconstriction in patients with hyper-reactive airways.  
50 Sequence 172 BP; 0 A; 35 C; 42 G; 39 T.

[illegible]

Search completed: Wed Jun 24 20:41:32 1998  
Job time : 740 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

\*\*\*\*\*

Run on: Wed Jun 24 19:02:10 1998; Masparr time 2788.51 Seconds  
1230.107 Million cell updates/sec

\*\*\*\*\*

\*\*\*\*\*  
Title: >US-08-866-354-1  
Description: (1-2568) from US08866354.seq  
Perfect Score: 2525  
N.A. Sequence: 1 AATCTGCGCTCGACACGC.....CACTCTGTACAAAGAAA 2568  
Comp: TTAGACGGAGCCTGTGTGCG.....GTAGAACATGTTTCTTT

Scoring table: TABLE default

Gap 6

Match STD : Dbase 0; Query 0

Searched: 1759237 seqs, 667866413 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

emb1-est54

1:em-est1 2:em-est3

Database:

genbank-est106

3:gb-est1 4:gb-est10 5:gb-est11 6:gb-est12 7:gb-est13

8:gb-est14 9:gb-est15 10:gb-est16 11:gb-est17

12:gb-est18 13:gb-est19 14:gb-est20

15:gb-est21 16:gb-est22 17:gb-est23 18:gb-est3 19:gb-est4 20:gb-est5

21:gb-est6 22:gb-est7 23:gb-est8 24:gb-est9 25:gb-est10

26:gb-est11

Statistics: Mean 12.402; Variance 3.221; scale 3.850

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.
C	1	495	19.6	515	7	AA512935	nh91b04.s1 NCI-CGAP_Br	0.00e+00
C	2	428	17.0	631	7	AA518362	vt10a02.r1 Barstead mo	0.00e+00
C	3	354	14.0	446	5	AA387098	vc18b07.r1 ko mouse em	0.00e+00
C	4	351	13.9	465	21	W53498	md45e01.r1 Soares mous	0.00e+00
C	5	330	13.1	430	21	W76716	ms81h06.r1 Soares mous	0.00e+00
C	6	303	12.0	330	9	AA573164	nm51f06.s1 NCI-CGAP_Br	0.00e+00
C	7	274	10.9	369	8	AA505897	n101c09.s1 NCI-CGAP_Br	0.00e+00
C	8	201	8.0	266	7	AA471183	pmw2139.KGla lambda za	3.39e-254
C	9	177	7.0	478	24	AA250120	mw21g08.r1 Soares mous	1.21e-216
C	10	158	6.3	467	13	AA780215	et52i08.s1 Soares tota	3.32e-187
C	11	118	4.7	278	4	AA12082	EST182778 Jurkat T-cel	3.32e-126
C	12	115	4.6	521	18	H12981	y170a10.r1 Homo sapien	1.04e-121
C	13	97	3.8	439	12	AA727460	vu96a10.r1 Stratagene	4.65e-95

14	94	3.7	229	3	R02249	ye83h05.r1 Homo sapien	1.12e-90	
15	82	3.2	459	24	AA245689	mx03e07.r1 Soares mous	2.24e-73	
C	16	252	2.5	13	AA754459	97SN1787 Rice Immature	5.19e-51	
C	17	62	2.5	13	AA754459	97SN1787 Rice Immature	1.38e-45	
C	18	51	2.0	13	AA754458	97SN1784 Rice Immature	3.90e-31	
C	19	47	1.9	13	AA754458	97SN1784 Rice Immature	4.30e-26	
C	20	43	1.7	23	AA064450	ml47h11.r1 Stratagene	3.39e-21	
C	21	42	1.7	350	4	AA349976	EST56965 Infant Brain	5.35e-10
C	22	38	1.5	2275	12	AF034173	Homo sapiens ntcon2 co	2.53e-15
C	23	36	1.4	284	25	FR0012397	F.rubripes GSS sequenc	4.58e-13
C	24	35	1.4	478	22	W73681	z53h01.r1 Soares feta	5.85e-12
C	25	30	1.2	318	9	AA587886	nm71a08.s1 NCI-CGAP.Co	1.09e-06
C	26	30	1.2	350	3	T09902	4455m3 Plasmodium falc	1.09e-06
C	27	30	1.2	360	25	B24410	F20A167R IGF Arabidops	1.09e-06
C	28	30	1.2	434	12	AA083088	zn08h06.s1 Stratagene	1.09e-06
C	29	30	1.2	444	15	AA819896	UT-R-A0-aq-c-09-0-UT.s	1.09e-06
C	30	30	1.2	466	9	AA524108	ng35a11.s1 NCI-CGAP.Co	1.09e-06
C	31	30	1.2	524	10	AA427938	zw53d03.s1 Soares tota	1.09e-06
C	32	30	1.2	565	10	AA610066	af08e07.s1 Soares test	1.09e-06
C	33	31	1.2	2275	12	AF034173	Homo sapiens ntcon2 co	1.05e-07
C	34	29	1.1	298	26	HS83467B5	H.sapiens (D1S2781) DN	1.07e-05
C	35	29	1.1	334	26	G10329	human STS CHLC.ATC6F11	1.07e-05
C	36	28	1.1	416	18	W78296	EST00444 Homo sapiens	1.00e-04
C	37	29	1.1	432	3	T47812	yb17c05.r1 Homo sapien	1.07e-05
C	38	28	1.1	518	25	B67798	T24H127R TAMU Arabidop	1.00e-04
C	39	28	1.1	522	7	C25513	D1cyosfeli1um discoidae	1.00e-04
C	40	29	1.1	655	25	B52556	CIT-HSP-2005L3.TR CIT-	1.07e-05
C	41	29	1.1	777	25	AG004298	Homo sapiens genomic D	1.07e-05
C	42	28	1.1	789	25	AG004298	Homo sapiens genomic D	1.00e-04
C	43	29	1.1	808	25	AG004297	Homo sapiens genomic D	1.07e-05
C	44	29	1.1	814	25	AG004330	Homo sapiens genomic D	1.07e-05
C	45	29	1.1	853	25	B19213	F10D23-T7 IGF Arabidop	1.07e-05

#### ALIGNMENTS

RESULT	1	AA512935	515 bp	mRNA	EST	05-AUG-1997
LOCUS		nh91b04.s1 NCI-CGAP_Br.1		Homo sapiens	CDNA	IMAGE:965839,
DEFINITION		mRNA sequence.				
ACCESSION		AA512935				
NID		92251358				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryota; Mitocondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;				
		Homo.				
REFERENCE		1 (bases 1 to 515)				
AUTHORS		NCI-CGAP http://www.nci.nih.gov/ncicgap.				
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
JOURNAL		Tumor Gene Index				
COMMENT		Unpublished (1997)				

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/BLND at:

www.bio.lnl.gov/btrp/image/image.html

Insert Length: 1735 Std Error: 0.00

Seq primer: -40m13 fwd. RT from Amersham

High quality sequence stop: 460.

Location/Qualifiers

1. 515

/organism="Homo sapiens"

/note="Vector: pRT3D-Pac (Pharmacia) with a modified





Y	1471	CAISGTGTGACTGCACGCAACAGTGGGAACGACTGAGAGAGCTTGAATTTTGGATT	1530
Db	434	TTTTTAAGGCAATACGTGTCTCAAAAATGCAATTTCAAGCTTTGGCAATGGCTCGATG	493
Oy	1531	TCTTCAAGACAAATACATGTCTTAAATAATGCAATTCAGAGCTTTGGCAATGGCTCGATG	1590
Db	494	TGACCATGTGGCAGCAGCAGCAGC-CACAGTCCAGACACCAATCCAGCTACCATGCT	552
Oy	1591	TGACCGTGTGGCAGCAGCAGCTTCCAGTACACACACCACTCCACTCCACACTGCC	1650
Db	553	TCCGATCAAGAACGAGCTACTGGG-CCGGCAGGTTCTGAGATAGAT	600
Oy	1651	TCCGGGTTAAGAACAGCCCTTGGGGCCACAGGCTGTGAGATGAAT	1699
RESULT	3		
LOCUS	AA387098	446 bp	EST
DEFINITION	vc18b07.r1	Ko mouse embryo 11 5dpc	Mus musculus cDNA clone 774901
ACCESSION	AA387098		
VERSIONS	g2040052		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryotes: mitochondrial eukaryotes; Metazoa; Chordata;		
AUTHORS	Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;		
	Murinae; Mus.		
	1 (bases 1 to 446)		
	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,		
	Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,		
	Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,		
	Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and		
	Waterston, R.		
TITLE	The WashU-HMNI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT			
	Contact: Marra M/Mouse EST Project		
	WashU-HMNI Mouse EST Project		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: mouseest@watson.wustl.edu		
	This clone is available royalty-free through LNL; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	MG1:467757		
	High quality sequence stop: 413.		
FEATURES			
source	Location/Qualifiers		
	1..446		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/note="Vector: pSPORT1; Site1: SalI; Site2: NotI; Total		
	RNAs were extracted from 11.5 dpc embryos (excluding		
	placenta and yolk sac). The double-stranded cDNA was		
	synthesized with an oligo (dT)-1 primer		
	GAGGAGACATGATCTTACATCGCGGCGCGCTTTTTTTTTTTTTTTT 3'.		
	The cDNAs were ligated to Lr-Sal3a. 5'		
	GCTATGACGTCGACTATCC 3' and Lr-Sal3b: 5'		
	GGAATGTCGAGCTCAAT 3'. The cDNAs were size-selected and		
	amplified by long-range PCR using Ex Taq polymerase for 18		
	cycles. The PCR-amplifiable cDNA mixture went through		
	one round of equalization and was digested with SalI/NotI		
	and cloned into the SalI/NotI sites of the pSPORT1		
	plasmid vector (Life Technologies). The library was		
	constructed by Dr. Minoru S. H. Ko and Dr. Xiaohong		
	Wang."		
	/db_xref="taxon:10090"		
	/clone="774901"		
	/clone_1b="Ko mouse embryo 11 5dpc"		
	/sex="pooled"		
	/dev_stage="11.5dpc"		
	/lab_host="DH10B"		
	<1..>446		
mRNA			

BASE COUNT	112 a	122 c	107 g	105 t
ORIGIN				
Query Match	14.0%; Score 354; DB 5; Length 446;			
Best Local Similarity	89.7%; Pred. No. 0.00e+00;			
Matches 400; Conservative	0; Mismatches 46; Indels 0; Gaps 0;			
Db	1 GCAGTCTTCGCACAAAGTTCCAGCCAGCCACACAGCTACGGAGTGTCTCTCTCTCCG 60			
Qy	1130 GCAGTCTTCCTTTCACAGGTCGCCGCCACACAGCTACGAGTGTCTCTCTCTCCG 1189			
Db	61 GGAGCTGCCCTGCACCCGAGAGGGCGGACACATATCTCTCTGTGTCTCTATGAGA 120			
Qy	1190 GGAGCTGCCCTGCACAGAGCGGAGCGGACACACATCTGTCTCTGTCTCTATGAGA 1249			
Db	121 ACGAGAGAGGCCCAACTCTCTGAATCTCAGAGTCTCTCAGAACAAATTCATCTGCA 180			
Qy	1250 GAGGAGAGAGGCCCAACTCTTGAATTCAGAGAGTCTCTCAGAACAAATTCATCTGCA 1309			
Db	181 ATCTGCGCTTGACATTTTCTTACCAGTCCAGCCAGACAGTCAAGTCTGTACAGACTG 240			
Qy	1310 ATCTGCGCTTGCGGATTTTCTTACCAGTCCAGCCAGACAGTCAAGTCTGTACAGACTG 1369			
Db	241 TCTTAAGAGAACTCGAGAGTGGCTCTGCGCCATCGGAGACTGATGGACAGTAT 300			
Qy	1370 TCTTAAGAGAACTCGAGTGGCTCTGCGCCATCGGAGACTGATGGACAGTAT 1429			
Db	301 GACTCTTAACATACATAGACTCCAGCAGCCTCAGTGTGCGCGTGTGTCGATTCAGCA 360			
Qy	1430 GACCCCAACATACATAGACTCCAGTGTGCGCGTGTGTCGATTCAGCA 1489			
Db	361 CAGTGGCAATGACCTGGAAGATTCGCTGAAGTTTCTGAATTTTAAAGACATACGTG 420			
Qy	1450 CAGTGGCAACGACCTGGAAGTGTGTAATTTTGAATTTCTTCAAGACATACATG 1549			
Db	421 TCTCAAAATGCAATTCAGCCTTG 446			
Qy	1550 TCTTAAATGCAATTCAGCCTTG 1575			
RESULT	4			
LOCUS	W53498 465 bp mRNA EST 03-JUN-1996			
DEFINITION	md5e01.t1 Soares mouse embryo NM013.5 14.5 Mus musculus cDNA clone 371352 5', mRNA sequence.			
ACCESSION	W53498			
NID	91357323			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	1 (bases 1 to 465)			
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucada,T., Lacy,M., Le,M., Martin,U., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.			
TITLE	The Washu-HHMI Mouse EST Project			
JOURNAL	Unpublished (1996)			
COMMENT	Contact: Marra M/Mouse EST Project Washu-HHMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:232784 Seq primer: mob.BEGA+ET High quality sequence stop: 347.			



QY 1510 AGCTCTGAATTTTGTGATTTCTTCAAGACAATACATGCTTAAATAATCAATTCAG 1569

Db 303 CCTTGCAGATGCTCGGATGTGACATGTGACAGCCGCCCCCGATCCAGACCA 362

QY 1570 CCTTGGCAATGACCTGATGACCGGTGGACAGCCTTCCAGTACAGACCA 1629

Db 363 CTGCGACACTACCTCCTCCGATCAAGAACACCTCTAGGCGCAGACGCTCTG 422

QY 1630 CTGCACACACACACCTCCCTCGGGTTAGAACACAGCCCTGGGCGCAGAGGCTGTG 1689

Db 423 AGAATGA 429

QY 1690 AGAATGA 1696

RESULT 6

LOCUS AA573164 330 bp mRNA EST 28-AUG-1997

DEFINITION nm51f06.s1 NCI\_CGAP\_Br2 Homo sapiens cDNA clone IMAGE:1071779, mRNA

ACCESSION AA573164

KEYWORDS 62347692

SOURCE EST.

ORGANISM human.

Homosapiens

Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 330)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.llnl.gov/bdrip/image/image.html

Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 319.

FEATURES

source

1. .330

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. This library is the normalized version of NCI-CGAP Br1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."

/db\_xref="taxon:9606"

/clone\_image="1071779"

/clone\_lib="NCI\_CGAP\_Br2"

/sex="female, pooled"

/tissue\_type="breast"

/lab\_host="DH10B"

<1. .>330

BASE COUNT 75 a 62 c 75 g 118 t

ORIGIN

MRNA

Query Match 12.0%; Score 303; DB 9; Length 330;

Best Local Similarity 99.4%; Pred. No. 0.00e+00;

Matches 310; Conservative 0; Mismatches 1; Gaps 1;

Db 20 TTTTTCCTTTTACATGTCATATGTAATTTTAAATGACGATCATGTTTCTGTA 79

CP 1980 TTTTTCCTTTTACATGTCATATGTAATTTTAAATGACGATCATGTTTCTGTA 1921

Db 80 AAGTAATAGGTGGAGACAGCGGTTACACAGACACACATGGGCTAGACACAGC 139

CP 1920 AAGTAATAGGTGGAGACAGCGGTTACACAGACACACATGGGCTAGACACAGC 1861

Db 140 TTGAGAGAGCAGCATGATTTGTGTTATGTGCTGAA-CACGAGACCTCTTTT 198

CP 1860 TTGAGAGAGCAGCATGATTTGTGTTATGTGCTGAA-CACGAGACCTCTTTT 1801

Db 199 CATATATACCATTTGAAATACAGAGGTGTATTTGCCGACATTTGATTCAGCTTCT 238

CP 1800 CATATATACCATTTGAAATACAGAGGTGTATTTGCCGACATTTGATTCAGCTTCT 1741

Db 259 GTCCCTTAATTTGACACAGGTGGCAAAACATGAGTGGAAATTTCTTCAGACCTGT 318

CP 1740 GTCCCTTAATTTGACACAGGTGGCAAAACATGAGTGGAAATTTCTTCAGACCTGT 1681

Db 319 CTGCCCCCAGGG 330

CP 1680 CTGCCCCCAGGG 1669

RESULT 7

LOCUS AA505897 369 bp mRNA EST 18-AUG-1997

DEFINITION n10ic09.s1 NCI\_CGAP\_Br2 Homo sapiens cDNA clone IMAGE:966736, mRNA

ACCESSION AA505897

NID 92242034

KEYWORDS EST.

SOURCE human.

ORGANISM human.

Homosapiens

Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 369)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.llnl.gov/bdrip/image/image.html

Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 338.

FEATURES

source

1. .369

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. This library is the normalized version of NCI-CGAP Br1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."

/db\_xref="taxon:9606"

/clone\_image="966736"

/clone\_lib="NCI\_CGAP\_Br2"

/sex="female, pooled"

/tissue\_type="breast"















soluble derivatives (extracellular domain), chimeric GDNF alpha and antibodies which bind to the GDNF alpha, including agonist and neutralising antibodies, as well as various uses for these molecules. It also relates to assay systems for detecting ligands to GDNF alpha, systems for studying the physiological role of GDNF, diagnostic techniques for identifying GDNF-related conditions, methods for identifying molecules homologous to GDNF alpha, and therapeutic techniques (claimed) for the treatment of GDNF-related and GDNF alpha-related conditions, particularly kidney disease associated with glomerulonephritis and enteric nervous system related disorders. Transgenic and knockout animals are also claimed.

Sequence 468 AA:

Query Match 94.5%; Score 3184; DB 27; Length 468;  
Best Local Similarity 92.9%; Pred. No. 0.00e+00;  
Matches 430; Conservative 27; Mismatches 5; Indels 1; Gaps 1;

1 mflctlyfajpjdllmsaevsgdrldcvkaadqklkegscstkyrtlrgcvaqketnf 60  
1 mflctlyfajpjdllmsaevsgdrldcvkaadqklkegscstkyrtlrgcvaqketnf 60  
1 mflctlyfajpjdllmsaevsgdrldcvkaadqklkegscstkyrtlrgcvaqketnf 60  
61 slsaglaekdecrrsmaealkqkslyncrcrkymkkeknclywmyqslqgnlldsp 120  
61 slsaglaekdecrrsmaealkqkslyncrcrkymkkeknclywmyqslqgnlldsp 120  
121 yepvnerladlfravplstevfgvehstkgmncldaaekcnldctckyrseyltpect 180  
121 yepvnerladlfravplstevfgvehstkgmncldaaekcnldctckyrseyltpect 180  
121 yepvnerladlfravplstevfgvehstkgmncldaaekcnldctckyrseyltpect 180  
181 smenevncrrckhalkrfqfdkvpahksymglfcscrdiacterirgltlrvpvcyeeer 240  
181 smenevncrrckhalkrfqfdkvpahksymglfcscrdiacterirgltlrvpvcyeeer 240  
181 smenevncrrckhalkrfqfdkvpahksymglfcscrdiacterirgltlrvpvcyeeer 240  
241 pncslajdscktnyicrsrladffncqpsrsyvnckkenyadclaysgilgtvmtpn 300  
241 pncslajdscktnyicrsrladffncqpsrsyvnckkenyadclaysgilgtvmtpn 300  
241 pncslajdscktnyicrsrladffncqpsrsyvnckkenyadclaysgilgtvmtpn 300  
301 yvassajsvwpwcdcsngndledclfnlffntclhnaigafgsgavtmqppv 360  
301 yvassajsvwpwcdcsngndledclfnlffntclhnaigafgsgavtmqppv 360  
301 yvassajsvwpwcdcsngndledclfnlffntclhnaigafgsgavtmqppv 360  
361 qttatctatfrvknkpjgpaeseneipthvlppcanlqgaqklknsygtchlcslsd 420  
361 qttatctatfrvknkpjgpaeseneipthvlppcanlqgaqklknsygtchlcslsd 420  
361 qttatctatfrvknkpjgpaeseneipthvlppcanlqgaqklknsygtchlcslsd 420  
421 gkgdglagasshlttkmaappscslslpvlmltalaaalsvs 463  
421 gkgdglagasshlttkmaappscslslpvlmltalaaalsvs 463  
421 gkgdglagasshlttkmaappscslslpvlmltalaaalsvs 463  
421 gkgdglagasshlttkmaappscslslpvlmltalaaalsvs 463

RESULT 2  
ID P80885 standard; protein: 505 AA.  
AC P80885;  
DT 16-DEC-1990 (first entry)  
DE Sequence encoded by LAV MA L GAG gene  
KM HIV: HTLV III: AIDS: diagnosis; vaccine; probe; hybridisation.  
OS lymphadenopathy associated virus MA L.  
PN W08707906-A.  
PD 30-DEC-1987.  
PR 22-JUN-1987; E00326.  
PR 23-JUN-1986; EP-401380.  
PA (INSP) Inst Pasteur.  
PI Alizon M, Sonigo P, Wain-Hobson S, Montagnier L;  
DR WPI; 88-014396/02.  
DR N-PSDB; N80437.  
PT New variants of lymphadenopathy associated virus (LAV) -  
PT used for prodn. of DNA, antigens and antibodies used in  
PT diagnosis of AIDS and pre-AIDS  
PS Claim 8: FIG 8a-8i: 72pp; English.  
CC LAV EL 1 (N80436) and LAV MA L (N80437) were isolated from the peripheral  
CC blood lymphocytes of patients. Different AIDS virus isolates concerned  
CC are designated by 3 letters of the patients name. Stable probes including

the DNA sequences can be used for detection of the new LAV viruses or  
CC related viruses or DNA proviruses in eg. biological samples. The proteins  
CC or peptides can be used for detection of antibodies induced in vivo and  
CC present in biological fluids. The DNA can also be used for the expression  
CC of LAV viral antigens for the prodn. of a vaccine against LAV. The  
CC polypeptides can also be used for the prodn. of antibodies for the  
CC detection of proteins related to the LAV viruses, partic. for diagnosis  
CC of AIDS or pre-AIDS.  
Sequence 505 AA:

Query Match 2.8%; Score 96; DB 1; Length 505;  
Best Local Similarity 41.9%; Pred. No. 1.61e+01;  
Matches 18; Conservative 11; Mismatches 9; Indels 5; Gaps 5;

28 kyr-lkhlywasrelerfalgplletgegcqqlmeql-qstl 68  
45 KYRTLRCQV-AGKETN-FSLASGL-EAKDECRSMAEALKOKSL 84

RESULT 3  
ID R15252 standard; protein: 32 AA.  
AC R15252;  
DT 12-FEB-1992 (first entry)  
DE Carbohydrate binding domain #9.  
KM cellulose; CBD; hemicellulosic substrate;  
KM Trichoderma reesei; cellulase; terminal A region.  
PN M09117244-A.  
PD 14-NOV-1991.  
PR 08-MAY-1991; DK0124.  
PR 09-MAY-1990; DK-001158.  
PA (NOVO) NOVO NORDISK A/S.  
PI Moldike HF, Hagen F, Hjort CM, Hastrup S;  
DR WPI; 91-353766/48.  
PT New fungal (hemi)cellulose degrading enzymes - for prodn. of 11q.  
PT fuel gas and feed protein, have specified carbohydrate binding domain  
PS Claim 20; Page 45; 73pp; English.  
CC This CBD is homologous to a terminal A region of Trichoderma reesei  
CC cellulase and effects binding of a protein to an insoluble  
CC cellulosic or hemicellulosic substrate. It is one of ten specific  
CC CBD's (see R15244-R15253) which correspond to the generic CBD  
CC formulae in R15242 and R15243. The CBD is incorporated into a fusion  
CC protein comprising a catalytic domain from a cellulase, e.g. A  
CC Bacillus endoglucanase, and optionally comprising a linking B domain  
CC from e.g. a fungal endoglucanase.  
Sequence 32 AA:

Query Match 2.6%; Score 87; DB 3; Length 32;  
Best Local Similarity 35.0%; Pred. No. 7.18e+01;  
Matches 7; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

14 tckspftck-kindfysgcq 32  
249 SCKTNYICRSRLADFTNCO 268

RESULT 4  
ID W43305 standard; protein: 119 AA.  
AC W43305;  
DT 01-APR-1998 (first entry)  
DE p17 antigen of human immunodeficiency virus type 1.  
KM Immunocassay; HIV-1; human immunodeficiency virus type 1; saliva;  
KM immobilised; p17 antigen; detection; nitrocellulose support; antibody.  
OS Human immunodeficiency virus type 1.  
PN US5695930-A.  
PD 09-DEC-1997.  
PR 05-AUG-1996; 692445.  
PR 10-NOV-1994; US-337670.  
PA (KILP/) KILPATRICK T J.  
PA (WEIN/) WEINSTEIN D E.  
PI KILPATRICK TJ, WEINSTEIN DE;  
DR WPI; 98-041291/04.  
PT Immunocassay for HIV-1 antibodies in saliva samples - using  
PT immobilised p17 antigen

PS Claim 1: Columns 11-14; 12pp; English.  
 CC The present sequence represents the p17 protein antigen of human  
 CC immunodeficiency virus type 1 (HIV-1). The invention relates to a method  
 CC for detecting antibodies to HIV-1, which comprises contacting saliva  
 CC from a fasting human with immobilised HIV-1 p17 antigen, preferably on a  
 CC nitrocellulose-containing solid support, and detecting any antigen-  
 CC antibody complex, optionally with a labelled second antibody. The method  
 CC can be used to detect exposure of a mammal, especially a human, to  
 CC HIV-1. Detecting anti-HIV antibodies in saliva is safer as HIV is not  
 CC transmitted via saliva, as opposed to blood. The assay is simple and  
 CC cheap to perform without laboratory equipment or trained personnel and  
 CC can be performed at home by the saliva donor.  
 SQ Sequence 119 AA;

Query Match 2.6%; Score 86; DB 27; Length 119;  
 Best Local Similarity 37.2%; Pred. No. 8.45e+01;  
 Matches 16; Conservative 12; Mismatches 10; Indels 5; Gaps 5;

Db 27 kxk-1khlvwasrelefaavnpjlletsegcrq1lq91-qps1 67  
 45 KYRTLROCV-AGKERN-FSLASGL-EANDECRSAMEALQKQSL 84

RESULT 5  
 ID P70272 standard; protein; 120 AA.  
 AC P70272;  
 DT 27-FEB-1991 (first entry)  
 DE The sequence encoding the HIV virus 14 KDa gag precursor  
 DE protein.  
 KW HIV virus; gag protein; AIDS; vaccine.  
 OS HIV virus.  
 PN EP-230222-A.  
 PD 29-JUL-1987;  
 PF 05-JAN-1987; 100064.  
 PR 06-JAN-1986; US-816645.  
 PA (HOFF) HOFFMANN-IA ROCHE AG.  
 PI Kramer R, Reddy P, Shaber M;  
 DR WPI: 87-207550/30.  
 DR N-PSDB: N70431.  
 PT Polypeptide obt'd. from HTLV-III gag-gene - used for detecting  
 PT AIDS virus and for prep'n. of vaccines giving protective immunity.  
 PS Claim 3; Page 13; 39pp; English.  
 CC The sequence encodes the HIV virus 14 KDa gag precursor gene.  
 CC The HIV virus gag protein or proteolytically derived products  
 CC may be used as a vaccine against HIV virus infection (AIDS).  
 CC See also N70427, N70428, N70429, N70430 and N70432.  
 SQ Sequence 120 AA;

Query Match 2.6%; Score 86; DB 2; Length 120;  
 Best Local Similarity 37.2%; Pred. No. 8.45e+01;  
 Matches 16; Conservative 12; Mismatches 10; Indels 5; Gaps 5;

Db 16 kxk-1khlvwasrelefaavnpjlletsegcrq1lq91-qps1 56  
 45 KYRTLROCV-AGKERN-FSLASGL-EANDECRSAMEALQKQSL 84

RESULT 6  
 ID R26326 standard; protein; 132 AA.  
 AC R26326;  
 DT 04-FEB-1993 (first entry)  
 DE HTLV-IIIB p17 sequence.  
 KW human immunodeficiency virus; HIV-1; HIV-2; glycoprotein; gag;  
 KW epitope mapping; monoclonal antibody; G11c1, G11h3; B4F8; DB3.  
 OS Human T-cell lymphotropic virus.  
 FH Key Location/Qualifiers  
 FT peptide 1..15  
 FT /label- #1  
 FT 11..25  
 FT /label- #2  
 FT 21..35  
 FT /label- #3  
 FT 31..46  
 FT peptide

FT /label- #4  
 FT peptide 41..55  
 FT /label- #5  
 FT peptide 51..65  
 FT /label- #6  
 FT peptide 62..78  
 FT /label- #7  
 FT peptide 73..86  
 FT /label- #8  
 FT peptide 78..92  
 FT /label- #9  
 FT peptide 86..115  
 FT /label- #9A  
 FT peptide 93..107  
 FT /label- #10  
 FT peptide 103..115  
 FT /label- #11  
 FT peptide 111..123  
 FT /label- #12  
 FT peptide 121..132  
 FT /label- #13

FT WO9212731-A.  
 PD 06-AUG-1992.  
 PF 22-JAN-1992; U00528.  
 PR 22-JAN-1991; US-644435.  
 PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.  
 PI Abraham P, Honnen WJ, Revesz K, Shang F;  
 DR WPI: 92-284421/34.  
 PT Pure glycoprotein contg. gag amino acid sequences - expressed on  
 PT surface of HIV-infected cells and monoclonal antibodies to the  
 PT glycoprotein, for immunosay and protecting against HIV infection  
 PS Claim 14; Fig 12; 51pp; English.  
 CC The synthetic peptides #1-13 were used in epitope mapping studies  
 CC of monoclonal antibodies against p17. ELISA plates were coated with  
 CC recombinant p17 (p17) and peptides #1-13 (representing the complete  
 CC sequence of HTLV-IIIB p17). Antibodies G11g1 and G11h3 (which react  
 CC with live HIV infected cells and recognise a glycoprotein contg. a  
 CC gag amino acid sequence and having mol. wt. 150,000 or 90,000,  
 CC respectively), bound strongly to p17 but did not bind to any of the  
 CC peptides tested. The claimed monoclonal antibodies which competitively  
 CC inhibit G11g1 or G11h3 are further defined by the fact that they do  
 CC not substantially bind to the peptides 1-13. The MAbs can be used  
 CC to determine the number of HIV-infected cells, to purify gag  
 CC antigens or can be coupled to a toxin for therapeutic use.  
 SQ Sequence 132 AA;

Query Match 2.6%; Score 86; DB 5; Length 132;  
 Best Local Similarity 37.2%; Pred. No. 8.45e+01;  
 Matches 16; Conservative 12; Mismatches 10; Indels 5; Gaps 5;

Db 28 kxk-1khlvwasrelefaavnpjlletsegcrq1lq91-qps1 68  
 45 KYRTLROCV-AGKERN-FSLASGL-EANDECRSAMEALQKQSL 84

RESULT 7  
 ID R55851 standard; Protein; 169 AA.  
 AC R55851;  
 DT 07-DEC-1994 (first entry)  
 DE HIV-1 MA delta-5-16 mutant.  
 KW HIV-1; matrix protein; MA; gene therapy; AIDS; deletion mutant.  
 OS Human immunodeficiency virus type 1.  
 PN WO9412513-A.  
 PD 09-JUN-1994.  
 PF 12-JAN-1993; U00267.  
 PR 23-NOV-1992; US-979966.  
 PA (HARD) HARVARD COLLEGE.  
 PI Essex ME, Lee TH, Yu X;  
 DR WPI: 94-200178/24.  
 PT Treating HIV infection with mutant HIV matrix polypeptide - pref.  
 PT produced by transfection with viral vector, inhibiting  
 PT replication, assembly or infectivity of virus particles  
 PS Disclosure; Fig. 1A-1F; 61pp; English.









\*\*\*\*\*  
 W I D E R E I F  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Bloocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MSrch\_pp protein - protein database search, using Smith-Waterman algorithm

on: Tue Jun 23 18:23:39 1998; Maspar time 23.53 Seconds  
 721.817 Million cell updates/sec  
 ular output not generated.

Title: >US-08-866-354-2  
 Description: (1-465) from US08866354.pep  
 Perfect Score: 3369  
 Sequence: 1 MFLATLYFALPLDLILSAE.....PLVLVYVATSLTSLTETS 465

Scoring table: PAM 150  
 Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: p1r56  
 1:p1r1 2:p1r2 3:p1r3 4:p1r4 5:n13d

Statistics: Mean 46.109; Variance 80.746; scale 0.571

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	171	5.1	24	2	S69080	glycosyl-phosphatidyl 3.50e-13
2	102	3.0	249	2	S09868	hypothetical protein 1.22e-01
3	101	3.0	352	2	S60024	bradykinin B1 recepto 1.68e-01
4	98	2.9	346	2	A70144	hypothetical protein 4.43e-01
5	93	2.8	169	2	I64089	protein-export protei 2.13e+00
6	95	2.8	336	2	D70030	alkanal monooxygenase 1.15e+00
7	94	2.8	411	2	H69158	LPS biosynthesis Rfbu 1.56e+00
8	93	2.8	600	2	S07638	spore coat protein SP 2.13e+00
9	95	2.8	706	2	S62933	hypothetical protein 1.15e+00
10	96	2.8	826	2	A60385	monocyte surface anti 8.37e-01
11	95	2.8	857	2	S33821	median body protein - 1.15e+00
12	94	2.8	1075	2	S54067	probable membrane pro 3.92e+00
13	91	2.7	475	2	S49886	probable membrane pro 1.56e+00
14	92	2.7	493	2	JC5621	epidermal growth fact 2.89e+00
15	92	2.7	563	1	VCWV7	env polypeptide - bab 2.89e+00
16	92	2.7	688	1	CIHUS	complement subcompone 2.89e+00
17	92	2.7	695	2	S05008	serine proteinase, ca 1.71e+01
18	86	2.6	120	5	ITAM	hiv-1 matrix protein 1.71e+01
19	86	2.6	167	2	C64361	hypothetical protein 1.71e+01
20	86	2.6	316	2	S58719	probable membrane pro 1.71e+01
21	89	2.6	327	2	A55356	urokinase-type plasmi 7.12e+00
22	87	2.6	342	2	S42885	beta-1,3-glucanase (E 1.28e+01
23	86	2.6	387	2	B49175	Motch A protein - mou 1.71e+01

24	86	2.6	478	1	FOVMVL	gag polyprotein - hum 1.71e+01
25	89	2.6	491	2	S49779	CD1 protein - yeast 7.12e+00
26	88	2.6	494	2	G64382	acetylactate synthase 9.56e+00
27	86	2.6	500	1	FOVM2	gag polyprotein - hum 1.71e+01
28	86	2.6	502	1	FOVM2	gag polyprotein - hum 1.71e+01
29	88	2.6	506	1	A38068	gag polyprotein - hum 9.56e+00
30	86	2.6	512	1	FOVMH3	gag polyprotein - hum 1.71e+01
31	88	2.6	521	1	FOVJST	gag polyprotein - hum 9.56e+00
32	88	2.6	521	2	S53091	gag polyprotein - hum 9.56e+00
33	87	2.6	521	2	S12152	gag polyprotein - hum 1.28e+01
34	88	2.6	522	1	FOVJG	gag polyprotein - hum 9.56e+00
35	87	2.6	642	2	G65371	acetyl-CoA synthetase 1.28e+01
36	86	2.6	646	2	S36586	E1 protein - human pa 1.71e+01
37	86	2.6	769	1	JC1121	leukocyte adhesion pr 7.12e+01
38	89	2.6	1033	2	S54506	probable membrane pro 1.28e+01
39	87	2.6	1146	2	B35962	protein-tyrosine kina 1.28e+01
40	87	2.6	1182	2	A35962	protein-tyrosine kina 1.28e+01
41	88	2.6	1403	2	S24548	homeotic protein pros 9.56e+00
42	88	2.6	1403	2	JQ1397	pros protein - fruit 9.56e+00
43	88	2.6	1407	2	A41089	neuronal precursor pr 9.56e+00
44	87	2.6	2471	2	A49128	cell-fate determining 1.28e+01
45	87	2.6	2531	2	S18188	notch protein homolog 1.28e+01

## ALIGNMENTS

Result	Entry	Title	Organism	Accessions	Reference	Authors	Journal	Title
1	569080	#type complete glycosyl-phosphatidylinositol-linked protein GDNFR-alpha - rat	formal_name Rattus norvegicus #common_name Norway rat #Feb-1998 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998	569080		Treanor, J.J.S.; Goodman, L.; de Sauvage, F.; Stone, D.M.; Poulsen, K.T.; Beck, C.D.; Gray, C.; Armanini, M.P.; Pollock, R.A.; Hefti, F.; Phillips, H.S.; Goddard, A.; Moore, M.W.; Bui-Bello, A.; Davies, A.M.; Asel, N.; Takahashi, M.; Vandlen, R.; Henderson, C.E.; Rosenthal, A.	Nature (1996) 382:80-83	Characterization of a multicomponent receptor for GDNF.
	569080	#status preliminary #molecule_type mRNA						
SUMMARY	##residues 1-24 ##label TRE	length 24	molecular-weight 2573	checksum 2584				
Query Match	5.1%; Score 171; DB 2; Length 24;							
Best Local Similarity	95.8%; Pred. No. 3.50e-13;							
Matches	23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;							
Db	1 MFLATLYFALPLDLILSAEVS6G 24							
Qy	1 MFLATLYFALPLDLILSAEVS6G 24							
RESULT	2							
ENTRY	S09868	#type complete						
TITLE	hypothetical protein UL103 - human cytomegalovirus (strain AD169)							
ORGANISM	#formal_name human cytomegalovirus, human herpesvirus 5							
DATE	07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Sep-1997							
ACCESSIONS	S09868							
REFERENCE	S09749							
Authors	Che, M.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; Hutchinson III, C.A.; Kouzarides, T.; Martignetti, J.A.; Peddie, E.; Satchwell, S.C.; Tomlinson, P.; Weston, K.M.; Barrett, B.G.							
Journal	Journal							
Title	Analysis of the protein-coding content of the sequence of							



```

TITLE      alkalal monooxygenase homolog yvbr - Bacillus subtilis
ORGANISM   #formal_name Bacillus subtilis
DATE       05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
26-Feb-1998

ACCESSIONS D70030
REFERENCE   A69380
AUTHORS     Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
            Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessières, P.;
            Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Bruns,
            A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
            Brusch, C.V.; Caldwell, B.; Capano, V.; Carter, N.M.;
            Choi, S.K.; Codani, J.J.; Connerthon, I.F.; Cummings, N.J.;
            Daniel, R.A.; Denizot, F.; Devigne, K.M.; Duessenhoef, A.;
            Enllich, S.D.; Emerson, P.T.; Ettlin, K.D.; Erington, J.;
            Fabret, C.; Ferrari, E.; Foulger, C.; Fritz, C.; Fujita,
            M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,
            S.Y.; Glaser, P.; Goffeau, A.; Gollightly, E.J.; Grandi, G.;
            Guisepi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
            C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
            Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
            Kasahara, Y.; Klier-blanchard, M.; Klein, C.; Kobayashi,
            Y.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.;
            Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
            Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
            Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
            M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
            M.; Ogawa, K.; Ogihara, A.; Oudea, B.; Park, S.H.; Parro,
            V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott,
            A.M.; Prescann, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;
            Rey, M.; Reynolds, S.; Rieger, M.; Rivalta, C.; Rocha, E.;
            Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.;
            Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
            Sekowska, A.; Seror, S.J.; Seror, P.; Shin, B.S.; Soldo,
            B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
            Takemaru, K.; Takeuchi, M.; Takamashi, A.; Tanaka, T.;
            Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
            Vandenbol, M.; Vannier, F.; Vassaret, A.; Viat, A.;
            Wambutt, R.; Wedler, E.; Wedler, H.; Weltzenegger, T.;
            Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasunoto,
            K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.;
            Yoshikawa, H.; Danchin, A.

#journal
#title      Nature (1997) 390:249-256
#accession  D70030
#status     preliminary; nucleic acid sequence not shown;
            translation not shown
#molecule_type DNA
#residues   1-336 #label KUN
#experimental_source strain 168

GENETICS
#gene       yvbr
#classification #superfamily yvbr protein
#summary     #length 336 #molecular_weight 37101 #checksum 2405

Query Match      2.8% Score 95; DB 2; Length 336;
Best Local Similarity 42.5%; Pred. NO. 1.15e+00;
Matches 17; Conservative 11; Mismatches 10; Indels 2; Gaps 2;

Db 132 NSGDEPPELELRYFKPSGNVNRQVRAIPGSDIVPTW 171
||||:| | | | | | | | | | | | | | | | | | | |
Oy 317 NSGNDLECLKRL-NPFDMNTCLKNAIQAF-GNGSDVTYW 354

RESULT 7
ENTRY   H69158 #type complete
TITLE   LPS biosynthesis Rfbv related protein - Methanobacterium
         thermocautotrophicum (strain Delta H)
ORGANISM #formal_name Methanobacterium thermocautotrophicum
DATE     05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
ACCESSIONS H69158
REFERENCE   A69000

#authors     Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.;
            Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.;
            Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.;
            Lumm, W.; Pothier, B.; Qiu, D.; Spedifora, R.; Vicarel, R.;
            Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivan, N.; Caruso,
            A.; Bush, D.; Saefer, H.; Patwell, D.; Pradhanakar, S.;
            McDougall, S.; Shimer, G.; Goyal, A.; Pietrovski, S.;
            Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling,
            J.; Reece, J.N.
            J. Bacteriol. (1997) 179:7135-7155
            Complete genome sequence of Methanobacterium
            thermocautotrophicum Delta H: functional analysis and
            comparative genomics.
#accession  H69158
#status     preliminary; nucleic acid sequence not shown;
            translation not shown
#molecule_type DNA
#residues   1-411 #label MTH
#cross-references GB:AE000666
#experimental_source strain Delta H

GENETICS
#gene       MTH450
#summary     #length 411 #molecular_weight 45966 #checksum 9796

Query Match      2.8% Score 94; DB 2; Length 411;
Best Local Similarity 24.3%; Pred. NO. 1.56e+00;
Matches 18; Conservative 26; Mismatches 25; Indels 5; Gaps 4;

Db 252 IESKDELDDVKICIIIGAPVDEEYRLRLGLASKSDVDFGPPVTEYPTWMAADV 311
||||:| | | | | | | | | | | | | | | | | | | |
Oy 76 MEALKOKSLCYCRCK-GMKKEKNCRLRYMSYSLQGN-DLLEDSPEYPSNRL--SDI 131

Db 312 F-VLPSELEGRLV 324
|||:| | | | | | | | | | | | | | | | | | | |
Oy 132 FRVYFISDVEQV 145

RESULT 8
ENTRY   S07638 #type complete
TITLE   spore coat protein Sp96 precursor - slime mold (Dictyostellium
         discoideum)
ORGANISM #formal_name Dictyostellium discoideum
DATE     07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
08-Sep-1997
ACCESSIONS S07638; A60942; B60942
REFERENCE   S07638
AUTHORS     Fosnaugh, K.L.; Loomis, W.F.
            Nucleic Acids Res. (1989) 17:9489
            Sequence of the Dictyostellium discoideum spore coat gene
            Sp96
#cross-references MUID:90067962
#accession  S07638
#molecule_type DNA
#residues   1-600 #label FOS
#cross-references EMBL:X16491; NID:g7373; PID:g295736
REFERENCE   A60942
AUTHORS     Tasaka, M.; Hasegawa, M.; Ozaki, T.; Iwabuchi, M.; Takeuchi,
            I.
            Cell Differ. Dev. (1990) 31:1-9
            Isolation and characterization of spore coat protein (sp96)
            gene of Dictyostellium discoideum.
#accession  A60942
#molecule_type DNA
#residues   1-155, 'T', 157-414, 'C', 416-600 #label TAS
#accession  B60942
#status     not compared with conceptual translation
#molecule_type mRNA
#residues   1-155, 'T', 157-414, 'C', 416-600 #label TA2
GENETICS
#introns    22/1
KEYWORDS     duplication; glycoprotein
FEATURE

```





Thu Jun 25 08:47:38 1998

US-08-866-354-2.rpt

Page 6

Search completed: Tue Jun 23 18:25:42 1998  
Job time : 123 secs.

---





QY 1 MFATLYFALPLDLLLSAEVSGGDRLDVCYKASDOCLKEQSCSTKYRTLRQCVAGKETNF 60  
DB 61 SLASGLEAKDECRSAMBALOKSLYNCRCRKGKKEKNCRLRIYWSMTOSLOGNDLLEDSP 120  
QY 61 SLASGLEAKDECRSAMBALOKSLYNCRCRKGKKEKNCRLRIYWSMTOSLOGNDLLEDSP 120  
DB 121 YEPAVNSRLSDIFRVRVPFISDVFOVEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180  
QY 121 YEPAVNSRLSDIFRVRVPFISDVFOVEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180  
DB 181 SVSNDVCNRRKCHKALROFEDKVPKHSYGMFCSCRDIACTERRQTIYVPCSYEEREX 240  
QY 181 SVSNDVCNRRKCHKALROFEDKVPKHSYGMFCSCRDIACTERRQTIYVPCSYEEREX 240  
DB 241 PNCILQDSCCTNYICRSRLADFTNCPESRSVSCCKENYADCLLAYSGLTGYWTPN 299  
QY 241 PNCILQDSCCTNYICRSRLADFTNCPESRSVSCCKENYADCLLAYSGLTGYWTPN 299  
DB 300 YIDSSLSVAPMCDSCNSGNDLECKLFNFEDKNTCLKNAIOAFNGSDVTWQAPVPV 359  
QY 301 YIDSSLSVAPMCDSCNSGNDLECKLFNFEDKNTCLKNAIOAFNGSDVTWQAPVPV 360  
DB 360 QTTATATTTALRVKKNRPLGPAAGSENEIPTHVLPCCANLQAKLSNVSGNTHLCISNGNY 419  
QY 361 QTTATATTTALRVKKNRPLGPAAGSENEIPTHVLPCCANLQAKLSNVSGNTHLCISNGNY 420  
DB 420 EKEGLGASSHITTKSMAAPSCGSLPLVLYVLTALSTLLSTLS 464  
QY 421 EKEGLGASSHITTKSMAAPSCGSLPLVLYVLTALSTLLSTLS 465

RESULT 2  
ID GDNF\_RAT STANDARD: PRT: 468 AA.  
AC Q62997;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE GDNF RECEPTOR ALPHA PRECURSOR (GDNF-ALPHA) (TGF-BETA RELATED  
DE NEUROTROPHIC FACTOR RECEPTOR 1).  
GN GDNFRA OR TRNRL.  
OS RATVUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-RETINA;  
RX MEDLINE: 96270513.  
JING S., WEN D., YU Y., HOLST P.L., LUO Y., FANG M., TAMIR R.,  
ANTONIO L., HU Z., CUPPLES R., LOUIS J.-C., HU S., ALTROCK B.W.,  
FOX G.M.;

CELL 85:1113-1124(1996).  
RL -1- FUNCTION: RECEPTOR FOR GDNF. MEDIATES THE GDNF-INDUCED  
CC AUTOPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR.  
CC -1- SUBUNIT: 2 MOLECULES OF GDNF-ALPHA ARE THOUGHT TO FORM A COMPLEX  
CC WITH THE DISULFIDE-LINKED GDNF DIMER AND WITH 2 MOLECULES OF RET.  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, BRAIN AND KIDNEY.  
CC -1- SIMILARITY: BELONGS TO THE GDNF FAMILY.  
DR EMBL: U59486; G1399863; -  
KW RECEPTOR; GLYCOPROTEIN; GPI-ANCHOR; MEMBRANE; SIGNAL.  
FT SIGNAL 1 24  
FT CHAIN 25 ?  
FT PROPEP 2 468  
FT DOMAIN 362 369  
FT CARBOHYD 59 59  
FT CARBOHYD 347 347  
FT CARBOHYD 406 406  
FT SEQUENCE 468 AA: 51649 MW: 6A7A2B2A CRC32:  
Query Match 94.5%; Score 3184; DB 1; Length 468;  
Best Local Similarity 92.9%; Pred. No. 0.00e+00;  
Matches 430; Conservative 27; Mismatches 5; Indels 1; Gaps 1;

DB 1 MFATLYFALPLDLLLSAEVSGGDRLDVCYKASDOCLKEQSCSTKYRTLRQCVAGKETNF 60  
QY 1 MFATLYFALPLDLLLSAEVSGGDRLDVCYKASDOCLKEQSCSTKYRTLRQCVAGKETNF 60  
DB 61 SLASGLEAKDECRSAMBALOKSLYNCRCRKGKKEKNCRLRIYWSMTOSLOGNDLLEDSP 120  
QY 61 SLASGLEAKDECRSAMBALOKSLYNCRCRKGKKEKNCRLRIYWSMTOSLOGNDLLEDSP 120  
DB 121 YEPAVNSRLSDIFRVRVPFISDVFOVEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180  
QY 121 YEPAVNSRLSDIFRVRVPFISDVFOVEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180  
DB 181 SVSNDVCNRRKCHKALROFEDKVPKHSYGMFCSCRDIACTERRQTIYVPCSYEEREX 240  
QY 181 SVSNDVCNRRKCHKALROFEDKVPKHSYGMFCSCRDIACTERRQTIYVPCSYEEREX 240  
DB 241 PNCILQDSCCTNYICRSRLADFTNCPESRSVSCCKENYADCLLAYSGLTGYWTPN 299  
QY 241 PNCILQDSCCTNYICRSRLADFTNCPESRSVSCCKENYADCLLAYSGLTGYWTPN 299  
DB 300 YIDSSLSVAPMCDSCNSGNDLECKLFNFEDKNTCLKNAIOAFNGSDVTWQAPVPV 359  
QY 301 YIDSSLSVAPMCDSCNSGNDLECKLFNFEDKNTCLKNAIOAFNGSDVTWQAPVPV 360  
DB 360 QTTATATTTALRVKKNRPLGPAAGSENEIPTHVLPCCANLQAKLSNVSGNTHLCISNGNY 419  
QY 361 QTTATATTTALRVKKNRPLGPAAGSENEIPTHVLPCCANLQAKLSNVSGNTHLCISNGNY 420  
DB 420 EKEGLGASSHITTKSMAAPSCGSLPLVLYVLTALSTLLSTLS 464  
QY 421 EKEGLGASSHITTKSMAAPSCGSLPLVLYVLTALSTLLSTLS 465

RESULT 3  
ID GDNF\_MOUSE STANDARD: PRT: 468 AA.  
AC P97785;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE GDNF RECEPTOR ALPHA PRECURSOR (GDNF-ALPHA) (TGF-BETA RELATED  
DE NEUROTROPHIC FACTOR RECEPTOR 1).  
GN GDNFRA OR TRNRL.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-DORSAL ROOT GANGLION;  
RX WATABE K.;  
RA SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDJ DATA BANKS.  
RL -1- FUNCTION: RECEPTOR FOR GDNF. MEDIATES THE GDNF-INDUCED  
CC AUTOPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR (BY  
CC SIMILARITY).  
CC -1- SUBUNIT: 2 MOLECULES OF GDNF-ALPHA ARE THOUGHT TO FORM A COMPLEX  
CC WITH THE DISULFIDE-LINKED GDNF DIMER AND WITH 2 MOLECULES OF RET  
CC (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (BY  
CC SIMILARITY).  
DR EMBL: AB000800; G1816442; -  
KW RECEPTOR; GLYCOPROTEIN; GPI-ANCHOR; MEMBRANE; SIGNAL.  
FT SIGNAL 1 24  
FT CHAIN 25 ?  
FT PROPEP 2 468  
FT DOMAIN 362 369  
FT CARBOHYD 59 59  
FT CARBOHYD 347 347  
FT CARBOHYD 406 406  
FT SEQUENCE 468 AA: 51782 MW: 6C64C182 CRC32:  
Query Match 94.5%; Score 3183; DB 1; Length 468;



FT	CARBOHYD	413	413	POTENTIAL.
SO	SEQUENCE	463 AA:	51598 MW:	0A2165C0 CRC32:
	Query Match	45.2%	Score 1522:	DB 1: Length 463:
	Best Local Similarity	51.3%	Pred. No. 0.00e+00:	
	Matches 203:	Conservative 91:	Mismatches 90:	Indels 12: Gaps 10:
D	37	QVDCRAVELCALAESNCSSRRITRLRCLAGADRN-TM---	LANK-ECQALAEVLEQSEPLY	91
Q	26	RLDCKKASDQCLKQSCSTKYRTLQCAKGETNFSLASGLSEADDECRSAEALAKOSLY		85
D	92	DCRCRGGKKELOCLQIOWSIHLGTECEEFYEAQSPYPTSRLSDFRLASIFSGAD		151
Q	86	NCRCRGGKKEKNCRLRITWYSKQSL-OGNDLLEDSPIEPVARSLSDFRYVPTFSDVFOQ		144
D	152	PVVSASNHCLDAKACNLANDNCKRLSSYSISICNREISPTERCNRRCKHALQPFEDRY		211
Q	145	VEHLPKGNCCIDAAKACNLDDICKKYSRATYITPCTTYS-NDVCNRRCKHALQPFEDRY		203
D	212	PSETTYRMLFSCQDQCAEARRQITLPSCEYEDKEKPNCLDLRLCRTDHLCSRLADF		271
Q	204	PAKSYGMLFSCQDQCAEARRQITLPSCEYEDKEKPNCLDLRLCRTDHLCSRLADF		263
D	272	HANCRASYRTTSPADYQACLSYAGMIGEDMTPNVYNSPRTGIYVSPKNCRRSGNM		331
Q	264	FTNCPQERSVSSCLKEKYADCLLAYSLTWTWPTNTIDS-S-LVAPPCDQSGNSND		321
D	332	EECEKFLDPTENPCFLNMAIQAFNGTQVNM-SPKQPTSAQAPVE-KTSPILPDDL		389
Q	322	LEECIKFLNFKDMTCLKNAIQAFNGSDVYVQPAFVQVQTATTTTALRYKKNKPLGPA		381
D	390	DSTS-LGTSYTTCTSTQSLKANNSELSMCTE		424
Q	382	GSENEIPTHVLPCCANLQAKLKNSVGNTHLCLSN		417
	RESULT	6		
ID	NRTR	HUMAN	STANDARD:	PRT: 464 AA.
AC	000451:			
DT	01-NOV-1997	(REL. 35, CREATED)		
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)		
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)		
DE	NEURTURIN RECEPTOR ALPHA PRECURSOR (NRNR-ALPHA) (NRNR-ALPHA) (TGF-BETA RELATED NEUROTROPHIC FACTOR RECEPTOR 2) (GDNF RECEPTOR BETA) (GDNFR-BETA).			
DE	GDNFRB OR TRNR2.			
OS	HOMO SAPIENS (HUMAN).			
	EUKARYOTA: METAQOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA: EUTHERIA: PRIMATES.			
	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 97325791.			
RA	BALOH R.H., TANSEY M.G., GOLDEN J.P., CREEDON D.J.,			
RA	HECKERROTH R.O., KECK C.L., ZIMONJIC D.B., POBESCU N.C.,			
RA	JOHNSON E.M., MILBRANDT J.:			
RL	NEURON 18:793-802(1997).			
CC	-1- FUNCTION: RECEPTOR FOR NEURTURIN. MEDIATES THE NRIN-INDUCED			
CC	AUTOPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR.			
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (BY			
CC	SMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE GDNF FAMILY.			
CC	EMBL: AF002700; G2145080; -.			
DR	MIM: 601956; -.			
KM	RECEPTOR: GLYCOPROTEIN; GPI-ANCHOR; MEMBRANE; SIGNAL.			
FT	SIGNAL	1	?	POTENTIAL.
FT	CHAIN	?	?	TGF-BETA RELATED NEUROTROPHIC FACTOR
FT	PROPEP	?	464	RECEPTOR 2.
FT	CARBOHYD	52	52	HYDROPHOBIC, REMOVED DURING MATURATION
FT	CARBOHYD	357	357	(POTENTIAL).
FT	CARBOHYD	413	413	POTENTIAL.
Q	SEQUENCE	464 AA:	51558 MW:	POTENTIAL. CRC32:

Query Match	45.28	Score 1522	DB 1	Length 464
Best Local Similarity	48.38	Pred. No. 0.00e+00		
Matches 223	Conservative 100	Mismatches 122	Indels 17	Gaps 1
Db	12	FLDETTLRLSPSSLOGPELHGWRPVDCVRANELCAAESNCSSRYRTLROCLAGDRN	70	
Oy	2	FLATLYPALDLDLLLSAEVSG - GDRLDVCVAKSDQCLKQSGSTYRTLROCVAGKETNF	60	
Db	71	TM--LANK-ECQALAEVLQESFLYDCRRKRGKKELQCLQYNSIHGLGEEFEYAS	126	
Oy	61	SLASGLEAMKECSAAEALOKSLYCNCRKRMKEKELCLRYMSYOSL -OGNDLLEDS	119	
Db	127	PYBPVTRLDIFRLMSIFSGTGADPVVSAKSHHCIDAKACNLNDNCKRISYSICN	186	
Oy	120	PYBPVNSRLSDIFRRVPFLISDVQOYEHPKGNCLDAKACNLNDICKKRSATITPCT	179	
Db	187	REISPTERCNRKCHKALROFEDRVPSXYTYRMLFCSCODACAERRQITLLPSCSYEDK	246	
Oy	180	TSVS -NDVCNRKCHKALROFEDVKPAKSYGMLFCSCDIACCTERRQITVPCSYER	238	
Db	247	EKPSPCDLRQVCTDHLCKSRDLAFANCRASVQYVTSPPADNYACIGSTAGMGFPMY	306	
Oy	239	EKPNCMLQOSCKTNYICRRLADFTNQPSRSVSCLEKNYADCLAYSGLLGYMT	298	
Db	307	PNVYDSPTGIYVSPMCSGSGSNMEDECEKFLRDPTEPCRNALQAFNGTDVNVSPK	366	
Oy	299	PNTYDS -S-LSTAPMCDGNSGNDLEELKFLNFKNQTKLAKNALQAFNGSDVTVQMP	356	
Db	367	GPSFOATQAPRV-E-KTPSLPDLSDSTIS-LQTSYITTCSTVSQEOGLKANNKSELSCFT	423	
Oy	357	AFPOQTATTATTTALVKKNPCLDPASGENIEPTHLVPCANLQAKLSNVSGNTHLCIS	416	
Db	424	ELTNTIIP-GSNKVIKPNSGPSARPSALTVLSYMLKQAL	464	
Oy	417	NGNYEKGDLGASHITTKS-MA-APPSGCLSPDLVLVVT-AL	455	
RESULT	7	STANDARD:	PRT:	465 AA.
AC	013157			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	NEURURIN RECEPTOR ALPHA PRECURSOR (NTNR-ALPHA) (GNDF			
DE	RECEPTOR BETA) (GNFR-BETA).			
GN	GNFRB.			
OS	GALLUS GALLUS (CHICKEN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;			
OC	GALLIFORMES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	BUI-BELLO A., ADU J., PINON L.G., HORTON A., THOMPSON J.,			
RA	ROSENTHAL A., CHINCHEIRO M., BUCHMAN V.L., DAVIES A.N.;			
RL	NATURE 387:721-724 (1997).			
CC	-1- FUNCTION: RECEPTOR FOR NEURURIN. MEDIATES THE NRTN-INDUCED			
CC	AUTOPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR.			
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (BY			
CC	SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE GNFR FAMILY.			
DR	EMBL, U90542; G2213805; -			
FT	RECEPTOR; GLYCOPROTEIN; GPI-ANCHOR; MEMBRANE; SIGNAL.			
FT	CHAIN 1 ? ?			
FT	PROPEP ? 465			
FT	CARBOHYD 355			
FT	CARBOHYD 387			
FT	CARBOHYD 412			
FT	SEQUENCE 465 AA; 51908 MW; 5CA073E4 CRC32;			
Query Match	43.98	Score 1480	DB 1	Length 465
Best Local Similarity	46.58	Pred No. 0.00e+00		



RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,  
 RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODER A., KELLEY J.M.,  
 RA WEIDMAN J.F., PHILLIPS R.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,  
 RA UTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUNDER D.M., BRANDON R.C.,  
 RA FINE L.D., FRITCHMAN J.L., FUHRMAN J.L., GEORGAGEN N.S.M.,  
 RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
 RA VENTER J.C.,  
 RA SCIENCE 269:496-512(1995).  
 CC -1- FUNCTION: THIS IS ONE OF THE PROTEINS REQUIRED FOR THE NORMAL  
 CC EXPORT OF ENVELOPE PROTEINS OUT OF THE CELL CYTOSOL. IT MAY  
 CC BE INVOLVED IN THE INITIATION OF THE EXPORTING PROCESS, BY BINDING  
 CC TO THE NASCENT POLYPEPTIDE VIA A SIGNAL SEQUENCE, MAINTAINING  
 CC A STABLE AND PRE-TRANSLATION CONFORMATION (BY SIMILARITY).  
 CC -1- SUBUNIT: ONE OF SEVEN SECRETORY PROTEINS (SECA-F & SECY) THAT  
 CC COMPRISE THE PROKARYOTIC PROTEIN TRANSLLOCATION APPARATUS  
 CC (BY SIMILARITY).  
 CC EMBL: U32158; G1573750; -.  
 DR TIGR: H10743; -.  
 RA PROTEIN TRANSPORT: TRANSLLOCATION.  
 CC SEQUENCE 169 AA: 19132 MW: 748840E CRC32;  
 Query Match 2.8%; Score 93; DB 1; Length 169;  
 Best Local Similarity 25.0%; Pred. No. 2.92e-01;  
 Matches 16; Conservative 17; Mismatches 30; Indels 1; Gaps 1;  
 Db 75 LEDSGDAFICEKQAGVFTISGLDVOAHCLTSCQPNMLFPAARELYSNLNGTFPA 134  
 Qy 246 LODSKNTVICSRLADFTFNQPSRSVSSCKENYADCLLAY-GLIGVTMTNYIDS 304  
 Db 135 LNL 138  
 Qy 305 SLS 308  
 RESULT 11 STANDARD; PRT; 504 AA.  
 ID GAG\_HYIMA  
 AC P04594;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
 DE GAG POLYPROTEIN (CONTAINS: CORE PROTEINS P17, P24, P2, P7, P1, P6).  
 GN GAG.  
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (MAL ISOLATE) (HIV-1).  
 OS VIRIDAE: SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;  
 CC LENTIVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 MEDLINE: 86245056.  
 ALIZON M., MAIN-HOBSON S., MONTAGNIER L., SONIGO P.;  
 CELL 46:63-74(1986).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FORM A VIRION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PFM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 DR EMBL: K03456; G328021; -.  
 DR EMBL: X04415; G60229; -.  
 DR EMBL: A07116; G492872; -.  
 DR HSSP: P03348; 1HVN.  
 DR HIV: K03456; GAGSMAL.  
 KW AIDS: CORE PROTEIN; POLYPROTEIN; MYRISTYLATION; PHOSPHORYLATION.  
 FT INT\_MET 0  
 FT LIPID 0  
 FT CHAIN 1 137 MYRISTATE (BY SIMILARITY).  
 FT CHAIN 138 368 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 369 383 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 384 437 CORE PROTEIN P2.  
 FT CHAIN 438 453 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 454 504 CORE PROTEIN P1.  
 FT CHAIN 504 504 CORE PROTEIN P6.  
 SO SEQUENCE 504 AA: 56001 MW: 4500CD44 CRC32;

Query Match 2.8%; Score 96; DB 1; Length 504;  
 Best Local Similarity 41.9%; Pred. No. 9.39e-02;  
 Matches 18; Conservative 11; Mismatches 9; Indels 5; Gaps 5;  
 Db 27 KYR-LKHLWASRELEPALNPGLLETGEGCOQIMEDL-QSTL 67  
 Qy 45 KYRLLROCV-AGKETN-FSLASGL-EAKDECRSMELAKQKSL 84  
 RESULT 12 STANDARD; PRT; 600 AA.  
 ID SP96.DICDI  
 AC P14328;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE SPORE COAT PROTEIN SP96.  
 GN COTA.  
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).  
 CC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;  
 CC EUMYCETOZOA; DICTYOSTELIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AX4;  
 RX MEDLINE: 90067962.  
 RA FOSNAUGH K., LOOMIS W.F.;  
 RL NUCLEIC ACIDS RES. 17:9489-9489(1989).  
 CC -1- SUBCELLULAR LOCATION: OUTER LAYER OF THE COAT MATRIX AND  
 CC INTERSPORE MATRIX.  
 CC -1- PFM: PHOSPHORYLATED AND FUCOSYLATED. MAY BE PHOSPHOGLYCOSYLATED,  
 CC MAY CONTAINS GLCNAC-ALPHA-1-P-SER RESIDUES.  
 CC -1- SIMILARITY: CONTAINS 4 PRESPORE MOTIFS.  
 DR EMBL: X16491; G295736; -.  
 DR PTR: S07638; S07638.  
 DR DICTYB: DD03007; COTA.  
 KW GLYCOPROTEIN; PHOSPHORYLATION; REPEAT; SPORULATION.  
 FT REPEAT 185 197 PRESPORE MOTIF 1.  
 FT REPEAT 221 233 PRESPORE MOTIF 2.  
 FT REPEAT 298 310 PRESPORE MOTIF 3.  
 FT REPEAT 395 407 PRESPORE MOTIF 4.  
 SO SEQUENCE 600 AA: 59589 MW: B18FCD86 CRC32;  
 Query Match 2.8%; Score 93; DB 1; Length 600;  
 Best Local Similarity 40.4%; Pred. No. 2.92e-01;  
 Matches 19; Conservative 8; Mismatches 14; Indels 6; Gaps 6;  
 Db 298 CRNIOCPFGICEDHNRNPICVLEBEPNDCLTCDNVCEASGLVC 344  
 Qy 216 CRDIAC-TE-R-RRQITVPVCSYEBERKP-NCLNLAD-SCRTN-YIC 256  
 RESULT 13 STANDARD; PRT; 706 AA.  
 ID YNCL\_YEAST  
 AC P53973;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 80.1 KD PROTEIN IN UME3-PUB1 INTERGENIC REGION.  
 GN YNLO21M OR N2819.  
 OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).  
 CC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ANDRE B., TIRAGUI ROUSSAINT I., URBESTARAZU L.A., VISSERS S.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SIMILARITY: BELONGS TO THE RPD3/C08B11.2/ACDC FAMILY.  
 DR EMBL: Z71297; E239664; -.  
 KW HYPOTHETICAL PROTEIN.  
 SO SEQUENCE 706 AA: 80069 MW: 04B8BCAD CRC32;  
 Query Match 2.8%; Score 95; DB 1; Length 706;  
 Best Local Similarity 26.6%; Pred. No. 1.38e-01;  
 Matches 21; Conservative 23; Mismatches 31; Indels 4; Gaps 3;

Db	330	LVIISGSDADMTIGCGHTPCYCIGHMILSLARGN--LCVVEGYNDIAIRSA	387
Oy	371	LIVANKRPLGPGRSENEIPTHYLPFCANLQAKKLSNVSGNTHLCIS-NGNIENEGGLASS	429
Db	388	LSVAKVLIGEPPDELDPDL	406
Oy	430	HITTKSMAPSPSCGLS-PL	447
RESULT	14		
ID	MS2_MOUSE	STANDARD;	PRT; 826 AA.
AC	005910;		
DT	01-NOV-1995 (REL. 32, CREATED)		
DI	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)		
DE	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)		
DE	CELL SURFACE ANTIGEN MS2 PRECURSOR (EC 3.4.24.-) (MACROPHAGE CYSTEINE-RICH GLYCOPROTEIN).		
GN	ADAM8 OR MS2.		
	MUS MUSCULUS (MOUSE).		
	EUKARYOTA; METAFAA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
	EUTHERIA; RODENTIA.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ICR:		
RA	YAMAMOTO S., YOSHIIYAMA K., SETOGUCHI M., MATSURA K., HIGUCHI Y., AKIZUKI S.;		
RL	SUBMITTED (JAN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.		
RN	[2]		
RP	PRELIMINARY SEQUENCE FROM N.A.		
RC	STRAIN-ICR:		
RA	YOSHIDA S., SETOGUCHI M., HIGUCHI Y., AKIZUKI S., YAMAMOTO S.;		
RL	INT. IMMUNOL. 2:585-591(1990).		
CC	-1- FUNCTION: POSSIBLE INVOLVEMENT IN EXTRAVASATION OF LEUCOCYTES.		
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
CC	-1- TISSUE SPECIFICITY: MACROPHAGES.		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B (ZINC METALLOPROTEASE); ALSO KNOWN AS THE REPROLYSIN SUBFAMILY.		
DR	EMBL: X13335. E264744; -		
DR	EMBL: X13335. E225615; -		
DR	MED: MG1:107825; ADAM8.		
DR	PROSITE: PS00142; ZINC_PROTEASE; 1.		
DR	PROSITE: PS00022; EGF_1; UNKNOWN.1.		
KW	PROSITE: PS01186; EGF_2; UNKNOWN.1.		
DR	TRANSMEMBRANE: GLYCOPROTEIN; ANTIGEN; ZINC; HYDROLASE;		
KM	METALLOPROTEASE; SIGNAL.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	17	826
FT	DOMAIN	17	658
FT	TRANSMEM	659	683
FT	DOMAIN	684	826
FT	METAL	329	329
FT	ACT_SITE	330	330
FT	METAL	333	333
FT	METAL	339	339
FT	CARBOHYD	89	89
FT	CARBOHYD	260	260
FT	CARBOHYD	431	431
FT	CARBOHYD	614	614
SO	SEQUENCE	826 AA;	90046 MW; 7C26F36F CRC32;
Query Match		2.8%;	Score 96; DB 1; Length 826;
Best Local Similarity		33.3%;	Pred. No. 9.39e-02;
Matches	17;	Conservative	10; Mismatches 18; Indels 6; Gaps 6;
Db	585	ELVLOGTNC-BEGLVC-MDGSCODLR-VYRSENCASCNHGVCHNRRECH	632
Oy	146	EHIPGNNCLDAKACNLDDICKKYRSAYITP-CITSVSND-YCN-BRKCH	193
RESULT	15		
ID	MEDB_GIALA	STANDARD;	PRT; 857 AA.

```

AC      Q08014;
DT      01-NOV-1995 (REL. 32, CREATED)
DT      01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT      01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE      MEDIAN BODY PROTEIN.
OS      GIARDIA LAMBLLA (GIARDIA INTESTINALIS).
OC      EUKARYOTA; PROTOZOA; SARCOMASTICOPHORA; NASTIGOPHORA; DIPLOMONADIDA;
OC      HEXAMITIDAE.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PORTLAND-1;
RX      MEDLINE; 93287123.
RA      MARSHALL J., HOLBERTON D.V.;
RL      J. MOL. BIOL. 231:521-530(1993).
CC      -I- FUNCTION: MAY HAVE A ROLE IN IMMOBILISING THE MICROTUBULES
CC      BETWEEN CELL DIVISIONS.
CC      -I- SUBCELLULAR LOCATION: MEDIAN BODY.
CC      -I- DOMAIN: SHOWS AN ALPHA-HELICAL COILED COIL STRUCTURE (30
CC      REPEATING HEPTADS).
DR      EMBL; X64517; G312671; -.
KW      CYTOSKELETON; MICROTUBULES; COILED COIL; HEPTAD REPEAT PATTERN.
SQ      SEQUENCE 857 AA; 100583 MW; 5E2BBAB4 CRC32;

Query Match          2.8%; Score 95; DB 1; Length 857;
Best Local Similarity 33.3%; Pred. No. 1.38e-01;
Matches 15; Conservative 16; Mismatches 11; Indels 3; Gaps 3.

Db      482 KENSYNFDDLEQKOQRSDNALREKAAYE-RVDSELRUKDE 525
        ||::: ||::: ||:::||:::| |:::|:::|:::|:::|:::|
Yy      56 KETNPSLASGLEAKDECRSMAMELKQSL-YNCCKRGMK-KEN 98
```

Search completed: Tue Jun 23 18:20:47 1998  
Job time : 68 secs.

THIS PAGE BLANK (USPTO)







DB 416 GKDGLAGASSHTTTSMAAPSCGSLSPVMTALALLSYS 458  
 ID 035748 PRELIMINARY: PRT: 463 AA.  
 AC 035748  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.  
 OS RATUS NORVEGICUS (RAT).  
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:  
 RN EUTHERIA: RODENTIA.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HANNOVER:  
 RA SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 EMBL: AJ002072; E1154274; -  
 SEQUENCE 463 AA: 51032 MM; 93277F91 CRC32;

Query Match 91.7%; Score 3091; DB 10; Length 463;  
 Best Local Similarity 91.8%; Pred. No. 0.00e+00;  
 Matches 425; Conservative 27; Mismatches 5; Indels 6; Gaps 2;

DB 1 MFLATYFALPLDLILMSAEVSGGDRDCVKASDCLKEQSCSTKYRTLROCVAKETNF 60  
 ID 035748 PRELIMINARY: PRT: 463 AA.  
 AC 035748  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.  
 OS RATUS NORVEGICUS (RAT).  
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:  
 RN EUTHERIA: RODENTIA.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HANNOVER:  
 RA SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 EMBL: AJ002072; E1154274; -  
 SEQUENCE 463 AA: 51032 MM; 93277F91 CRC32;

DB 1 MFLATYFALPLDLILMSAEVSGGDRDCVKASDCLKEQSCSTKYRTLROCVAKETNF 60  
 ID 035748 PRELIMINARY: PRT: 463 AA.  
 AC 035748  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.  
 OS RATUS NORVEGICUS (RAT).  
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:  
 RN EUTHERIA: RODENTIA.  
 RN (1)  
 RP SEQUENCE FROM N.A.

RA TRUPP M., RAYNOSCHER C., IBANEZ C.F.;  
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 ID 035748 PRELIMINARY: PRT: 463 AA.  
 AC 035748  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.  
 OS RATUS NORVEGICUS (RAT).  
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:  
 RN EUTHERIA: RODENTIA.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HANNOVER:  
 RA SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 EMBL: AJ002072; E1154274; -  
 SEQUENCE 463 AA: 51032 MM; 93277F91 CRC32;

Query Match 45.3%; Score 1525; DB 10; Length 464;  
 Best Local Similarity 49.6%; Pred. No. 0.00e+00;  
 Matches 209; Conservative 93; Mismatches 106; Indels 13; Gaps 11;

DB 12 FLDETSLASPSLSQSELGMRPOVDCVANELCAESNCSRYRTLROCLAGDRN- 70  
 ID 035748 PRELIMINARY: PRT: 463 AA.  
 AC 035748  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.  
 OS RATUS NORVEGICUS (RAT).  
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:  
 RN EUTHERIA: RODENTIA.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HANNOVER:  
 RA SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 EMBL: AJ002072; E1154274; -  
 SEQUENCE 463 AA: 51032 MM; 93277F91 CRC32;

DB 12 FLDETSLASPSLSQSELGMRPOVDCVANELCAESNCSRYRTLROCLAGDRN- 70  
 ID 035748 PRELIMINARY: PRT: 463 AA.  
 AC 035748  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.  
 OS RATUS NORVEGICUS (RAT).  
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:  
 RN EUTHERIA: RODENTIA.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HANNOVER:  
 RA SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 EMBL: AJ002072; E1154274; -  
 SEQUENCE 463 AA: 51032 MM; 93277F91 CRC32;



OY 236 EREKPNCLNLDSCRTNCRSLADFTNCPESRSVSSCLKENYADCLAYSLIST 295

DB 236 AMTPNFISVNTTVALSTCRGSGNLQDECEQLERSFSQNPCLVEAIAA 344  
 OY 236 VMTPNYIDSSLSVAPWCDSCNSGNDLECLKFLNFKNTCLKNAIOA 344

RESULT 9  
 ID 035325 PRELIMINARY: PRT: 397 AA.

AC 035325:  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR FAMILY RECEPTOR ALPHA-3.  
 GN GFR-3.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:  
 RN EUTHERIA: RODENTIA.

SEQUENCE FROM N.A.  
 TRUPP M., RAYNOSCHER C., IBANEZ C.F.  
 SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF020305; G2429367;  
 SQ SEQUENCE 397 AA; 44333 MW; F0C0C841 CRC32;

Query Match 21.2%; Score 714; DB 10; Length 397;  
 Best Local Similarity 35.2%; Pred. No. 4,75e-136;  
 Matches 123; Conservative 70; Mismatches 135; Indels 21; Gaps 16;

DB 11 LLMILLVLSLMLPLGAGSLATENFVSCQARKECANPACKAAYOHLGSCSTSLR 70  
 OY 2 FLATLYFALPL-LDLLLSAEVSGDRL-D-CYKASDQCLEKSCSKRYTLRCQVAGKET 58  
 DB 71 PLPLEES-AMSDCLLEAABQLRNSLIDCRHRMRKQATCLDIYTVHPARSLGDELD 129  
 OY 59 NLSLSGLEAKDECRSAMALOKSLYNCRCRKMKEKNCRLIYMSYQSLQ-GNDLLE 117  
 DB 130 VSPYE-DVYTSKPKM-N-LSYL-NMK--PDSCLLFAMCTLDHDCDLRKRKYGGA 182  
 OY 118 DSEYEVNRSRLDIFRVNPFISDVQVEHIFPGNNCCLAAKCNLDICIKRYRXYTP 177

DB 183 CS-GIR---CORHLCLADRSFEFEKAESHAOGLILCPCPEDAGGGERINTAPSCAL 238  
 OY 178 CITSVSNVNCNRKCHKALRQFDKYPAKHSYGMFLCSCR-DIACCTERRQIIVVCY 235  
 DB 239 PS-VTPNCLDLRSFCRADPLCRSLMDFOTHCPMDI-LGTCATEQ-SRCLRAYIGLIGT 295  
 OY 236 EREKPNCLNLDSCRTNCRSLADFTNCPESRSVSSCLKENYADCLAYSLIST 295  
 DB 236 AMTPNFISVNTTVALSTCRGSGNLQDECEQLERSFSQNPCLVEAIAA 344  
 OY 296 VMTPNYIDSSLSVAPWCDSCNSGNDLECLKFLNFKNTCLKNAIOA 344

RESULT 10  
 ID 080624 PRELIMINARY: PRT: 492 AA.

AC 080624:  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)  
 DE GAG PROTEIN.

GN GAG.  
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).  
 OC VIRIDAE: SS-RNA ENVELOPED VIRUSES: POSITIVE-STRAND: RETROVIRIDAE:  
 RN LENTIVIRINAE.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VI205; TISSUE-BLOOD; BRENNAN T., PETERS M., BRENNAN T.,  
 RA LOWMAGE J.J., MCCUTCHAN F., BRENNAN T., PETERS M., BRENNAN T.,  
 RA SANDERS-BUELL E., EDDY G., DER GROEN G., FRANSEN K.,  
 RA GERHRY-DAMET M., DELEYS R., BURKE D.,  
 RA SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: L11785; G555036;  
 SQ SEQUENCE 492 AA; 54981 MW; 930293D1 CRC32;

Query Match 3.1%; Score 103; DB 11; Length 492;  
 Best Local Similarity 42.5%; Pred. No. 8,13e-02;  
 Matches 17; Conservative 9; Mismatches 10; Indels 4; Gaps 4;

DB 26 SKKYR-LKHIWASRELKRFALNPLGLETSBGCROIIOL 64  
 OY 43 STKRTLRCV-AGKE-TNFSLASGL-EAKDECRSAMEAL 79

RESULT 11  
 ID 097845 PRELIMINARY: PRT: 129 AA.

AC 097845:  
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)  
 DE MATRIX PROTEIN P17 (FRAGMENT).

GN GAG.  
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).  
 OC VIRIDAE: SS-RNA ENVELOPED VIRUSES: POSITIVE-STRAND: RETROVIRIDAE:  
 RN LENTIVIRINAE.

SEQUENCE FROM N.A.  
 RP STRAIN-BAB MNT;  
 RC NARA R., ROUES P., COURROTIN C., PARNETMATHIEU F., BOUSSIN F.,  
 RA ROANE A., MARC D., LASFARGUES G., DORMONT D.,  
 RA J. VIROL. 70:4474-4483(1996).  
 DR EMBL: X99948; E258593;  
 KW MATRIX PROTEIN.

FT NON\_TER 129  
 SQ SEQUENCE 129 AA; 14550 MW; 542AB924 CRC32;

Query Match 3.0%; Score 100; DB 11; Length 129;  
 Best Local Similarity 41.9%; Pred. No. 2,23e-01;  
 Matches 18; Conservative 11; Mismatches 9; Indels 5; Gaps 5;

DB 28 KYR-LKHIWASRELKRFALNPLGLETSBGCROIIOL-QSTL 68  
 OY 45 KYRTLRCV-AGKETN-FSLASGL-EAKDECRSAMEALKOKSL 84

RESULT 12  
 ID 036786 PRELIMINARY: PRT: 130 AA.

AC 036786:  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
 DE MA-P17 (FRAGMENT).

GN GAG.  
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).  
 OC VIRIDAE: SS-RNA ENVELOPED VIRUSES: POSITIVE-STRAND: RETROVIRIDAE:  
 RN LENTIVIRINAE.

RP SEQUENCE FROM N.A.  
 RA LEIGH BROWN A.J., LOBIDEL D., MADE C.M., REBUS S., PHILLIPS N.,  
 RA BRETTLE R.P., FRANCE A.J., LEEN C.S., WOMENAMIN J., MONTILLAN A.,  
 RA MAW R.D., MULCAHY F., ROBERTSON J.R., SANKAR K.N., SCOTT G., WYLD R.,  
 RA PEUTHERER J.F.,  
 RA VIROLOGY 235:166-177(1997).  
 DR EMBL: AF014183; G2406708;  
 FT NON\_TER 130  
 FT NON\_TER 130  
 SQ SEQUENCE 130 AA; 14584 MW; 428D9E75 CRC32;

Query Match 3.0%; Score 100; DB 11; Length 130;  
 Best Local Similarity 41.9%; Pred. No. 2,23e-01;  
 Matches 18; Conservative 11; Mismatches 9; Indels 5; Gaps 5;

DB 5 KYR-LKHIWASRELKRFALNPLGLETSBGCROIIOL-QPSL 45  
 OY 45 KYRTLRCV-AGKETN-FSLASGL-EAKDECRSAMEALKOKSL 84

```

RESULT 13
ID 036808 PRELIMINARY; PRT: 130 AA.
AC 036808;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE MA-P17 (FRAGMENT).
GN GAG.
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
LN LENTIVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RA LEIGH BROWN A.J., LOBIDEL D., MADE C.M., REBUS S., PHILLIPS N.,
RA BETTLE R.P., FRANCE A.J., LEEN C.S., MCMENAMIN J., MCMILLAN A.,
RA MAM R.D., MULLACHY F., ROBERTSON J.R., SANKAR K.N., SCOTT G., WILD R.,
RA PEUTHERER J.F.;
VIRIOLOGY 235:166-177(1997).
EMBL: AF014205; G2406752; -.
FT NON_TER 130 130
FT NON_TER 1 1
SQ SEQUENCE 130 AA; 14527 MW; 7325723C CRC32;

Query Match 3.0%; Score 100; DB 11; Length 130;
Best Local Similarity 39.5%; Pred. No. 2,23e-01;
Matches 17; Conservative 12; Mismatches 9; Indels 5; Gaps 5;

Db 5 KYR-LKHIWASRELERFALNPGILETSEGCQIQMEQL-QSAL 45
OY 45 KYRTRROCV-AGKETN-FSLASGL-EAKDECRSAMEALKOKSL 84

RESULT 14
ID 097725 PRELIMINARY; PRT: 133 AA.
AC 097725;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DE 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE MATRIX PROTEIN P17 (FRAGMENT).
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
LN LENTIVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-PAL-W.
RA NARWA R., ROQUES P., COURPOTIN C., PARNETMATHIEU F., BOUSSIN F.,
RA ROANE A., MARC D., LASFARGUES G., DORMONT D.;
J. VIROL. 70:4474-4483(1996).
EMBL: Z79562; E261908; -.
FT NON_TER 133 133
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 14724 MW; E22E819C CRC32;

Query Match 3.0%; Score 100; DB 11; Length 133;
Best Local Similarity 44.2%; Pred. No. 2,23e-01;
Matches 19; Conservative 9; Mismatches 10; Indels 5; Gaps 5;

Db 28 KYR-LKHIWASRELERFALNPGILETSEGCQIQMEQL-QPSL 68
OY 45 KYRTRROCV-AGKETN-FSLASGL-EAKDECRSAMEALKOKSL 84

RESULT 15
ID 077804 PRELIMINARY; PRT: 486 AA.
AC 077804;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE GAG PROTEIN.
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;

```

```

OC LENTIVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-DJ258; TISSUE-BLOOD;
RA LOUWAGIE J.J., MCCUTCHAN F., BRENNAN T., PETERS M., BRENNAN T.,
RA SANDERS-BUELL E., EDDY G., DER GROEN G., FRANSEN K.,
RA GERSHY-DAMET M., DELEYS R., BURKE D.;
RL SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL: L11763; G534842; -.
SQ SEQUENCE 486 AA; 54094 MW; 6CB6088C CRC32;

Query Match 3.0%; Score 100; DB 11; Length 486;
Best Local Similarity 32.9%; Pred. No. 2,23e-01;
Matches 24; Conservative 18; Mismatches 24; Indels 7; Gaps 7;

Db 28 KYR-LKHIWASRELERFALNPGILETSEGCQIQMEQL-QSAL-GTESK-ELKSLYNTIA 83
OY 45 KYRTRROCV-AGKETN-FSLASGL-EAKDECRSAMEALKOKSLYNCRCKRGMKKKNCCLR 101
Db 84 VLMCVHQRIDIKD 96
OY 102 IYMSMYOSLOGND 114

Search completed: Tue Jun 23 18:23:21 1998
Job time : 137 secs.

```

\*\*\*\*\*  
 [W] [O] [R] [E] [H]  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Jun 24 20:41:51 1998: Maspar time 2627.46 Seconds  
 1357.834 Million cell updates/sec

\*\*\*\*\*  
 Similar output not generated.  
 \*\*\*\*\*

Title: >US-08-866-354-3  
 Description: (1-2138) from US08866354.seq  
 Perfect Score: 2138  
 N.A. Sequence: 1 AGCTGCGTCTCCCGGGGCGAG.....AGACAGCGCCGACGCCGTCG 2138  
 Comp: TCGACCGAGAGGGGCCCGTC.....TCGTGCGGGCGTCGCGCAGC

Scoring table: TABLE default  
 Gap 6

Rmatch STD : Dbase 0: Query 0

Searched: 457396 seqs, 834342348 bases x 2  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database:

emb154  
 1:em\_db 2:em\_htg 3:em\_hum1 4:em\_hum2 5:em\_in 6:em\_om  
 7:em\_ov 8:em\_pat 10:em\_pi 11:em\_ro 12:em\_v1  
 genbank106  
 13:gb\_db 14:gb\_htg 15:gb\_in 16:gb\_om 17:gb\_ov 18:gb\_pat  
 19:gb\_ph 20:gb\_pi 21:gb\_pi1 22:gb\_pi2 23:gb\_ro 24:gb\_st  
 25:gb\_sy 26:gb\_un 27:gb\_v1

Statistics: Mean 11.685; Variance 5.512; scale 2.120

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2138	100.0	2138	RNU59486	Rattus norvegicus GDNF	0.00e+00
2	2091	97.3	3616	RNU97142	Rattus norvegicus RET	0.00e+00
3	1615	75.5	2549	AF014117	Mus musculus GDNF rece	0.00e+00
4	1379	64.5	1551	AF015172	Mus musculus GDNF rece	0.00e+00
5	1302	60.9	1392	RN2072	Rattus norvegicus mRNA	0.00e+00
6	1279	59.8	1415	AB000800	Mouse mRNA for GDNF re	0.00e+00
7	1245	58.2	2560	AF038421	Homo sapiens GPI-1like	0.00e+00
8	1139	53.3	1619	AF042080	Homo sapiens glial cel	0.00e+00
9	1063	51.1	2175	HSU95847	Human GDNF receptor al	0.00e+00
10	1063	51.1	1707	HSU97144	Homo sapiens RET ligand	0.00e+00
11	660	29.9	777	AF012811	Mus musculus GDNF rece	0.00e+00
12	578	27.0	3037	GGU90541	Gallus gallus GDNF rec	0.00e+00
13	275	12.9	596	HSGFRA1G06	Homo sapiens GPI-1like	1.19e-205
14	244	11.4	497	HSGFRA1G03	Homo sapiens GPI-1like	2.41e-178
15	225	10.5	1395	HSU93703	Human glial cell line	1.09e-161

16	225	10.5	1526	AF002700	Homo sapiens GDNF fam1	1.09e-161
17	225	10.5	2906	HSU97145	Homo sapiens RET ligand	1.09e-161
18	220	10.3	1392	AF002701	Mus musculus GDNF fam1	2.57e-157
19	215	10.1	1395	AF005226	Rattus norvegicus glia	5.98e-153
20	213	10.0	2787	RN097143	Rattus norvegicus RET	3.32e-151
21	210	9.4	2933	GGU90542	Gallus gallus neururi	1.47e-122
22	166	7.8	529	HSGFRA1G02	Homo sapiens GPI-1like	1.67e-110
23	145	6.8	432	HSGFRA1G09	Homo sapiens GPI-1like	1.46e-92
24	114	5.3	832	HSGFRA1G11	Homo sapiens GPI-1like	1.70e-66
25	112	5.2	385	AA2902	Sequence 34 from paten	7.76e-65
26	106	5.0	397	HSGFRA1G08	Homo sapiens GPI-1like	7.04e-60
27	98	4.6	436	HSGFRA1G07	Homo sapiens GPI-1like	2.55e-53
28	65	3.0	363	HSGFRA1G04	Homo sapiens GPI-1like	3.78e-27
29	59	2.8	7218	HSGFRA1G04	Sequence 15 from paten	1.29e-22
30	42	2.0	215	182878	Sequence 5 from paten	1.97e-10
31	43	2.0	256	HSGFRA1G10	Homo sapiens GPI-1like	4.11e-11
32	42	2.0	7218	186494	Sequence 14 from paten	1.97e-10
33	33	1.7	215	182878	Sequence 5 from paten	1.79e-06
34	29	1.4	2602	S66545	putative alkaline phos	3.25e-02
35	31	1.4	129402	HS402611	Human DNA sequence ***	2.21e-03
36	29	1.4	134226	THICG	Ictalurid herpesvirus	3.25e-02
37	28	1.3	390	S65019	muscin (rats, Sprague-D	1.30e-01
38	27	1.3	565	E04076	gDNA encoding envelope	4.32e-01
39	28	1.3	1203	AF051767	Homo sapiens GDNF fam1	1.20e-01
40	28	1.3	10772	AF012089	Drosophila melanogaste	1.20e-01
41	28	1.3	134743	HUAC002300	Homo sapiens Chromosom	1.20e-01
42	28	1.3	143701	HS29C18	Homo sapiens DNA sequ	1.20e-01
43	28	1.3	165197	HSAC002070	Human BAC clone 7E1 f	1.20e-01
44	27	1.3	171368	HSU95738	Human chromosome 16p13	4.32e-01
45	28	1.3	202233	AC003664	Homo sapiens chromosome	1.20e-01

#### ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
1	RNU59486	Rattus norvegicus GDNF receptor alpha mRNA, complete cds.	U59486	g1399862	Norway rat, Rattus norvegicus	Eukaryotes; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Vertebrata	1 (bases 1 to 2138)	Jing, S.-J., Wen, D., Yu, Y., Holst, P. L., Luo, Y., Fang, M., Tamir, R., Antonio, L., Hu, Z., Cupples, R., Louis, J.-C., Hu, S., Altrock, B. and Fox, G. M.	GDNF-induced activation of the ret protein tyrosine kinase is mediated by GDNF-alpha, a novel receptor for GDNF	Cell 85 (7), 1113-1124 (1996)	/organism="Rattus norvegicus" /db_xref="taxon:10116" /codon_start=1 /product="GDNF receptor alpha" /translation="MPLATLYFALPLDILMSAFVSGDRDLCKVAKSDGCKECSST KYRTLRQCVAGKEINSESLTSGLEKDCRSAMBLKSLKSLNCRKGRKRCALIT TWSYVSLQGLDLEDSPIEYVNSRLDITRAVFIISDVROVYHISKGNCCDAARA CMLDDTCKYRSAAITPTCTSMSEVNRKRRCHALQFDDKVPKRSYGLNFCSCD IACERRRRTIVPCSEYERPNCLSLQSCSKNTYICRSRLADFFNCPSPHSVSN CLKENYADCLLAAYSGLTGVTWTPRYVDSLSVAPMCDSCNSGNDLEDCLEFLNFDD
							2 (bases 1 to 2138)	Fox, G. M., Jing, S.-J., Yu, Y., Holst, P. L., Fang, M., Tamir, R., Antonio, L., Hu, Z., and Louis, J.-C.	Direct Submission	Submitted (29-May-1996) Immunology, Amgen, Inc., Amgen Center, Thousand Oaks, CA 91320-1789, USA	Location/Qualifiers 1..2138 /organism="Rattus norvegicus" /db_xref="taxon:10116" 302..1708 /codon_start=1 /product="GDNF receptor alpha" /translation="MPLATLYFALPLDILMSAFVSGDRDLCKVAKSDGCKECSST KYRTLRQCVAGKEINSESLTSGLEKDCRSAMBLKSLKSLNCRKGRKRCALIT TWSYVSLQGLDLEDSPIEYVNSRLDITRAVFIISDVROVYHISKGNCCDAARA CMLDDTCKYRSAAITPTCTSMSEVNRKRRCHALQFDDKVPKRSYGLNFCSCD IACERRRRTIVPCSEYERPNCLSLQSCSKNTYICRSRLADFFNCPSPHSVSN CLKENYADCLLAAYSGLTGVTWTPRYVDSLSVAPMCDSCNSGNDLEDCLEFLNFDD



NTCLKNAIQFNGSDVTMMQAPVPVOTTTATTTAFRYVKNRPLGAGSENIPTHTL  
PCANLOAKLSNVSSTHLCISDSFGKDLGAGASSHITTTSMAPBSCSLSLPV  
LMITLAALLSVSLAERS"  
BASE COUNT 494 a 593 c 584 g 467 t  
ORIGIN

Query Match 100.0%; Score 2138; DB 23; Length 2138;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 2138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

D 1 AGCTCGCTCTCCGGGGCAGTGTGTGATGCACCGAGTCCGGCGCTGGGCAAGTTGG 60  
O 1 AGCTCGCTCTCCGGGGCAGTGTGTGATGCACCGAGTCCGGCGCTGGGCAAGTTGG 60  
D 61 GTGCGAATGAACCCCTGAAAGCGGGGTCCGCTCCCGCCCTCGGGCCCCCGGGATCTGA 120  
O 61 GTGCGAATGAACCCCTGAAAGCGGGGTCCGCTCCCGCCCTCGGGCCCCCGGGATCTGA 120  
D 121 GTGCGTGGCGCGGTGGGGCGGAGAGCGAGCGGGGAGTCTGCTTCACCCCTGATGAGCT 180  
O 121 GTGCGTGGCGCGGTGGGGCGGAGAGCGAGCGGGGAGTCTGCTTCACCCCTGATGAGCT 180  
D 181 GAACCTTGAGTGGCCAGAGAGCGCAGTCCGCCGGGATCGCTGCAGCGCTGACTCTCTC 240  
O 181 GAACCTTGAGTGGCCAGAGAGCGCAGTCCGCCGGGATCGCTGCAGCGCTGACTCTCTC 240  
D 241 CCCGAGACCGGGCGGGCTTTGATTTTGGGGGGGGGGGACCAAGCTGCGCGGGCGAC 300  
O 241 CCCGAGACCGGGCGGGCGGGCTTTGATTTTGGGGGGGGGGGACCAAGCTGCGCGGGCGAC 300  
D 301 CATGTTCTTACCCACTCTGTACTTCGCGGCCACTCTCTGATTTGCTGATGTCCGCGA 360  
O 301 CATGTTCTTACCCACTCTGTACTTCGCGGCCACTCTCTGATTTGCTGATGTCCGCGA 360  
D 361 GGTGAGTGTGTGAGAGACCGTCTGACTGTGTGAAAGCCAGCGCATGAGTCTGAAGAA 420  
O 361 GGTGAGTGTGTGAGAGACCGTCTGACTGTGTGAAAGCCAGCGCATGAGTCTGAAGAA 420  
D 421 GAGTGTGAGACCAAGTACCGCACACTAAGCAATGGCTGGCGGGGCAAGAAACCACTT 480  
O 421 GAGTGTGAGACCAAGTACCGCACACTAAGCAATGGCTGGCGGGGCAAGAAACCACTT 480  
D 481 CAGCTGACATCCGGCCCTTGAGGCCAAGAGATGATGCGGTAGCCCAATGAGAGCTTGA 540  
O 481 CAGCTGACATCCGGCCCTTGAGGCCAAGAGATGATGCGGTAGCCCAATGAGAGCTTGA 540  
D 541 GCAGAGTCTCTGTACAACTGCCCTGCAGAGCGGGGATGAAGAAAGAAATGTCT 600  
O 541 GCAGAGTCTCTGTACAACTGCCCTGCAGAGCGGGGATGAAGAAAGAAATGTCT 600  
D 601 GCGTATCTACTGGAGCATGTACAGAGCCCTGCAGGAAATGACCTCCTGGAAATTTCCC 660  
O 601 GCGTATCTACTGGAGCATGTACAGAGCCCTGCAGGAAATGACCTCCTGGAAATTTCCC 660  
D 661 GATAGACCCGGTTAAACAGAGTGTAGATATATTCGGGCACTCCCGTTTCAATACAGA 720  
O 661 GATAGACCCGGTTAAACAGAGTGTAGATATATTCGGGCACTCCCGTTTCAATACAGA 720  
D 721 TGTTTTCAGCAAGTGAACACATTTCCAAAGGAGAACACTGCTGAGCAGCAAGGC 780  
O 721 TGTTTTCAGCAAGTGAACACATTTCCAAAGGAGAACACTGCTGAGCAGCAAGGC 780  
D 781 CTGCAACCTGAGCAGACACTGTAGAGTACAGTCCGCCCTACACTACCCCTTGACACAC 840  
O 781 CTGCAACCTGAGCAGACACTGTAGAGTACAGTCCGCCCTACACTACCCCTTGACACAC 840  
D 841 CAGCATCTCCAACAGAGTCTGCAACCCCGTAAAGTGCACAAAGCCCTCAGAGCAATCTT 900  
O 841 CAGCATCTCCAACAGAGTCTGCAACCCCGTAAAGTGCACAAAGCCCTCAGAGCAATCTT 900  
D 901 CGACAAAGTTCGGGCCAAGACACACTACGGGATGCTTGTGCTCTGCGGGGACATCGC 960  
O 901 CGACAAAGTTCGGGCCAAGACACACTACGGGATGCTTGTGCTCTGCGGGGACATCGC 960

D 961 CTGCAACCGAGCGGGCGAGACACTATCTGCCCGCTGTGCTCTATGAAGAACGAGAG 1020  
O 961 CTGCAACCGAGCGGGCGAGACACTATCTGCCCGCTGTGCTCTATGAAGAACGAGAG 1020  
D 1021 GCCCAACTGCTGAGTGTGCAAGACTCTCAGAGCAATATACATGCGATCGGCT 1080  
O 1021 GCCCAACTGCTGAGTGTGCAAGACTCTCAGAGCAATATACATGCGATCGGCT 1080  
D 1081 TGCAGATTTTTTACCAACTGCGAGCCAGATCAAGTCTGTGAGCAACTGTCTTAAAGA 1140  
O 1081 TGCAGATTTTTTACCAACTGCGAGCCAGATCAAGTCTGTGAGCAACTGTCTTAAAGA 1140  
D 1141 GAACCTAGCAGACTGCTCTGCGCTACTCTGGGAGTATGGGACAGTCAATGACTCCCA 1200  
O 1141 GAACCTAGCAGACTGCTCTGCGCTACTCTGGGAGTATGGGACAGTCAATGACTCCCA 1200  
D 1201 CTACGTAGACTCCAGCAGCCTCAGCGTGGCACCATGGTGTGACTGAGCAACAGCGCAA 1260  
O 1201 CTACGTAGACTCCAGCAGCCTCAGCGTGGCACCATGGTGTGACTGAGCAACAGCGCAA 1260  
D 1261 TGACCTGGAAAGACTGCTTGAATTTTCTGAATTTTTTAAAGCAATACTTGTCAAAAA 1320  
O 1261 TGACCTGGAAAGACTGCTTGAATTTTCTGAATTTTTTAAAGCAATACTTGTCAAAAA 1320  
D 1321 TGCATTCAGCCCTTTGGCAATGGCTCAGATGTGACATGTGGCAGCAGCCCTCAGT 1380  
O 1321 TGCATTCAGCCCTTTGGCAATGGCTCAGATGTGACATGTGGCAGCAGCCCTCAGT 1380  
D 1381 CCAGACCCACCACTGCCACCACTACCACTGCCCTCCGGGTCMAAGCAAGCCTCTGGGGCC 1440  
O 1381 CCAGACCCACCACTGCCACCACTACCACTGCCCTCCGGGTCMAAGCAAGCCTCTGGGGCC 1440  
D 1441 AGCAGGCTCTGAAGATGAGATCCCCACACAGTTTAAACACCTGTGCGAATTTGCAAGC 1500  
O 1441 AGCAGGCTCTGAAGATGAGATCCCCACACAGTTTAAACACCTGTGCGAATTTGCAAGC 1500  
D 1501 TCAGAAAGCTAAATCCAAATGTGCGGGTACACACACCTGTGCTTTCGATGTGATTT 1560  
O 1501 TCAGAAAGCTAAATCCAAATGTGCGGGTACACACACCTGTGCTTTCGATGTGATTT 1560  
D 1561 CGGAAAGAGATGTCTGCTGCTGCTCCACAGCCACATTAACCAAAAAATCAATGGCTGCC 1620  
O 1561 CGGAAAGAGATGTCTGCTGCTGCTCCACAGCCACATTAACCAAAAAATCAATGGCTGCC 1620  
D 1621 TCCAGCTGCAAGTGTGAGTCACTGCGGGTGTGATGTCTACCGCCCTTGTCTCCTGTT 1680  
O 1621 TCCAGCTGCAAGTGTGAGTCACTGCGGGTGTGATGTCTACCGCCCTTGTCTCCTGTT 1680  
D 1681 ATCTGATCGTTGGCAAGAACGTGAGTGCATCCGGGGGAAACAGATGAAAAAGCAAA 1740  
O 1681 ATCTGATCGTTGGCAAGAACGTGAGTGCATCCGGGGGAAACAGATGAAAAAGCAAA 1740  
D 1741 AGAGAACCAAGTATTCGTCTGCTGCTCTGTATATCTGAAATCCAGTTTAAAGCT 1800  
O 1741 AGAGAACCAAGTATTCGTCTGCTGCTCTGTATATCTGAAATCCAGTTTAAAGCT 1800  
D 1801 CCGTTGAGAAAGCTTTACCCCACTGGAACCTTTTCTGTTTAAAGAAAGCTTGTGG 1860  
O 1801 CCGTTGAGAAAGCTTTACCCCACTGGAACCTTTTCTGTTTAAAGAAAGCTTGTGG 1860  
D 1861 CCGTCAAGGGGCTCTGTGTAAGAACTGTACAGGGCTAATTCCAAAACCAATAAGGGCTGG 1920  
O 1861 CCGTCAAGGGGCTCTGTGTAAGAACTGTACAGGGCTAATTCCAAAACCAATAAGGGCTGG 1920  
D 1921 GGGCGTGTGCGGCTTAAAGGGACATTTGCAACATGTAAGCAAGCAAGCTGGCTTATCATG 1980  
O 1921 GGGCGTGTGCGGCTTAAAGGGACATTTGCAACATGTAAGCAAGCAAGCTGGCTTATCATG 1980  
D 1981 TGTTTGATGATGAGATGAGTGTGATGTAGATGTGTAATTTTAAACAGCTTGAACCTTG 2040  
O 1981 TGTTTGATGATGAGATGAGTGTGATGTAGATGTGTAATTTTAAACAGCTTGAACCTTG 2040

Db 2041 TTCCTCTACTGTTAGAGACAGAGATACATATGATAAGATCTTCATGCTTACTC 2100  
|||||  
Oy 2041 TTCCTCTACTGTTAGAGACAGAGATACATATGATAAGATCTTCATGCTTACTC 2100  
Db 2101 AGCAGCATTCCTTGTGAAGACAGCGCCGACCGCTCG 2138  
|||||  
Oy 2101 AGCAGCATTCCTTGTGAAGACAGCGCCGACCGCTCG 2138  
RESULT 2 RN097142 3616 bp mRNA ROD 29-JUL-1997  
LOCUS Rattus norvegicus RET ligand 1 (RET1) mRNA, complete cds.  
DEFINITION U97142  
ACCESSION 92282021  
NID  
KEYWORDS  
SOURCE Norway rat.  
ORGANISM Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;  
Murinae; Rattus.  
REFERENCE 1 (bases 1 to 3616)  
Santicola, M., Hession, C. A., Worley, D. S., Carmillo, P., Ehrenfels, C.,  
Walus, L., Robinson, S., Jaworski, G., Wei, H., Tizard, R., Whitty, A.,  
Pepinsky, R. B. and Cate, R. L.  
Glial cell line-derived neurotrophic factor-dependent RET  
activation can be mediated by two different cell-surface accessory  
proteins  
Proc. Natl. Acad. Sci. U.S.A. 94 (12), 6238-6243 (1997)  
JOURNAL 9732356  
MEDLINE 2 (bases 1 to 3616)  
REFERENCE Santicola, M., Hession, C. A., Worley, D. S., Carmillo, P., Ehrenfels, C.,  
Walus, L., Robinson, S., Jaworski, G., Wei, H., Tizard, R., Whitty, A.,  
Pepinsky, R. B. and Cate, R. L.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-1997) Molecular Genetics, BIOGEN, 14 Cambridge  
Center, Cambridge, MA 02142, USA  
FEATURES  
source  
1..3616  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/dev\_stage="embryonic day 18"  
/tissue\_type="kidney"  
1..3616  
/gene="RET1"  
257..1863  
/gene="RET1"  
/note="PI-linked cell-surface accessory protein; GDNF  
receptor alpha"  
/codon\_start=1  
/product="RET ligand 1"  
/db\_xref="PID:92282022"  
/translation="MFLATLYFALPLDLMASAEVSGDRDLCVAKASDOLEKESCT  
KYRTLROCVAKETWFSLTSGLEAKDECRAMEALOKSLNCRKRMKREKELRI  
YMSYOSLOGNDLEDSDYEPVNSRLSDIFRAVPIISVFOOVAKHSYGLNCLDARA  
CNLDOTCKKYSAYTTPCTTSMNSNVNCRRCRCHALRQFMRPAKHSYGLNCLDARA  
IACRBRKQTIYVPCSYERERPNCLSLDSCKNRYIRSLADEFTTCOPESYSN  
CLKENYADCLAYSLGIVTWPNTVDSLSVAPWCSSNGNDLECLFLNFKD  
NCTKNAOAFGNSDVTMOPAPVQTTATTTTAFTVKKRPLGPAESEIEPTHL  
PPCANLQKLSKNSVSGSTHCLSDSDGKDLGASSHITTKMAAPPCSLSLPY  
LMLTALALLKLSVLAETS"  
BASE COUNT 876 a 955 c 921 g 864 t  
ORIGIN  
Query Match 97.3%; Score 2081; DB 23; Length 3616;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 2082; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 112 CGGATCTGAGTCGTGGCGGGGTGGCGGCGAGACGAGCGGAGTGTGCTCACCTG 171  
|||||  
Db 127 GATGAGCTGAACCTTTAGTGGCCAGAGAGAGGACAGTCCCGGGGATTCGTGACCTG 186  
|||||  
Oy 172 GATGAGCTGAACCTTTAGTGGCCAGAGAGAGGACAGTCCCGGGGATTCGTGACCTG 231  
|||||  
Db 187 AGCTCTCTCCCGAGACCGGGCGGGCTTGGATTTTGGGGGGGGGACACAGTCCG 246  
|||||  
Oy 232 AGCTCTCTCCCGAGACCGGGCGGGCTTGGATTTTGGGGGGGGGACACAGTCCG 291  
|||||  
Db 247 CGGGGACACATGTTCTTCTAGCCACTGTGTACTTGGCGCTGACCTCTGTATTTGCTGAT 306  
|||||  
Oy 292 CGGGGACACATGTTCTTCTAGCCACTGTGTACTTGGCGCTGACCTCTGTATTTGCTGAT 351  
|||||  
Db 307 GTCCGCGCAGTGAAGTGGTGAAGACCGTCTGTGACTGTGTGAAGACCATGAGTGGCT 366  
|||||  
Oy 352 GTCCGCGCAGTGAAGTGGTGAAGACCGTCTGTGACTGTGTGAAGACCATGAGTGGCT 411  
|||||  
Db 367 GAAGGACAGAGCTGACAGCAGCAGCAAGTACCGACACTAAGGCAAGTGGCGGCGCAAGGA 426  
|||||  
Oy 412 GAAGGACAGAGCTGACAGCAGCAGCAAGTACCGACACTAAGGCAAGTGGCGGCGCAAGGA 471  
|||||  
Db 427 AACCAACTTCAGCTGACATCCGGCCTTGAAGCCAAAGATGATGCCGTACGCCCATGGA 486  
|||||  
Oy 472 AACCAACTTCAGCTGACATCCGGCCTTGAAGCCAAAGATGATGCCGTACGCCCATGGA 531  
|||||  
Db 487 GGCCTTGAAGCAGAGTCTCTGTACACTGCGCTGCAAGCGGGGCGATGAAGAAAGAGA 546  
|||||  
Oy 532 GGCCTTGAAGCAGAGTCTCTGTACACTGCGCTGCAAGCGGGGCGATGAAGAAAGAGA 591  
|||||  
Db 547 GAATGTGTGGGTATCTCTGAGCATGTACAGAGCGCTGCAAGGAAATGACTCTCTGGA 606  
|||||  
Oy 552 GAATGTGTGGGTATCTCTGAGCATGTACAGAGCGCTGCAAGGAAATGACTCTCTGGA 651  
|||||  
Db 607 AGATTCCTCCGTATGAGCGCGGTTAAACAGCAGGTTGTCAATATATTCGCGGAGTCCCGTT 666  
|||||  
Oy 652 AGATTCCTCCGTATGAGCGCGGTTAAACAGCAGGTTGTCAATATATTCGCGGAGTCCCGTT 711  
|||||  
Db 667 CATATCAGATGTTTCCAGCAAGTGAACACATTTCCAAAGGAAACACTCCCTGAGCG 726  
|||||  
Oy 712 CATATCAGATGTTTCCAGCAAGTGAACACATTTCCAAAGGAAACACTCCCTGAGCG 771  
|||||  
Db 727 AGCCAAAGGCTGCAACCTGAGCAGACACTGTAAAGTACAGGTGGCGCTACATCACCC 786  
|||||  
Oy 772 AGCCAAAGGCTGCAACCTGAGCAGACACTGTAAAGTACAGGTGGCGCTACATCACCC 831  
|||||  
Db 787 CTGACACACAGCATGTCTCAACGAGGTGTGCAACCGCGTAAGTGCCAAGGCCCTCAG 846  
|||||  
Oy 832 CTGACACACAGCATGTCTCAACGAGGTGTGCAACCGCGTAAGTGCCAAGGCCCTCAG 891  
|||||  
Db 847 GCAGTCTTTCAGCAAGGTTCCGGCCAGACAGTACGAGGATGCTCTTCTCTCTCTCGCG 906  
|||||  
Oy 892 GCAGTCTTTCAGCAAGGTTCCGGCCAGACAGTACGAGGATGCTCTTCTCTCTCTCGCG 951  
|||||  
Db 907 GGAATGCGCTGCAACCGGAGCGGGCGGAGCACTATGTCTCCCGTGTCTCTATGAAGA 966  
|||||  
Oy 952 GGAATGCGCTGCAACCGGAGCGGGCGGAGCACTATGTCTCCCGTGTCTCTATGAAGA 1011  
|||||  
Db 967 AGAGAGAGGCGCCAACTCCCTGAGTCTGCAAGACTCTCTGCAAGCAATTCATCTGCG 1026  
|||||  
Oy 1012 AGAGAGAGGCGCCAACTCCCTGAGTCTGCAAGACTCTCTGCAAGCAATTCATCTGCG 1071  
|||||  
Db 1027 ATCTGCGCTTGCAGATTTTATTTTACCACTGCCAGCAGAGTCAAGTCTGACAGACTG 1086  
|||||  
Oy 1072 ATCTGCGCTTGCAGATTTTATTTTACCACTGCCAGCAGAGTCAAGTCTGACAGACTG 1131  
|||||  
Db 1087 TCTTAAGGAGAACTACGAGAGTGTCTCTGCGCTACCTGGGAGCTGATGGACAGTAT 1146  
|||||  
Oy 1132 TCTTAAGGAGAACTACGAGAGTGTCTCTGCGCTACCTGGGAGCTGATGGACAGTAT 1191  
|||||  
Db 1147 GACTCCCAACTACGATGAGACTCAGAGCTCAGGCTGACAGCTGATGATGATGATGACGAA 1206  
|||||

OY	1192	GACTCCCACTACGTATACCTTCAGACAGCCTCACGCGGCCACCATGGTGTGACTGCAGCAA	1251
Db	1207	CAGGGCAATGAGACTGGAAAGACGCTTGAAATTTCGATTTTTTTAAGACAATACCTGG	1266
OY	1352	CAGGGCAATGAGACTGGAAAGACGCTTGAAATTTCGATTTTTTTAAGACAATACCTGG	1311
Db	1367	TCTCAAAAATGCAATTTCAAGCTTTGGCAATGGCTCAGATGTACCATGTGGAGAGCAC	1326
OY	1312	TCTCAAAAATGCAATTTCAAGCTTTGGCAATGGCTCAGATGTACCATGTGGAGAGCAC	1371
Db	1327	CCCTCAGTCCAGACACCACTGCCACCACTACACTGCTTCCGGGTCAAGAACAAGCC	1386
OY	1372	CCCTCAGTCCAGACACCACTGCCACCACTACACTGCTTCCGGGTCAAGAACAAGCC	1431
Db	1387	TCTGGGGCCAGCAGGGGCTGTAGATAGATCCCCACACACGTTTTCACACCCTGTGCGAA	1446
OY	1432	TCTGGGGCCAGCAGGGGCTGTAGATAGATCCCCACACACGTTTTCACACCCTGTGCGAA	1491
	1447	TTTTCAGAGCTCAGAACGCTGAAATCCAAATGTGTGGGGTAGACACACACCTGTGCTTGTGA	1506
	1492	TTTTCAGAGCTCAGAACGCTGAAATCCAAATGTGTGGGGTAGACACACACCTGTGCTTGTGA	1551
Db	1507	TAGTGAATTTGCGAAAGAGATGCTCGCTGCTGCTGCCTCAGACACATTAACACAATAATCAAT	1566
OY	1552	TAGTGAATTTGCGAAAGAGATGCTCGCTGCTGCTGCCTCAGACACATTAACACAATAATCAAT	1611
Db	1567	GCGTGTCTCTCCACAGCTGACGTGACCTCACTGCGGTGTGTGATGATCACC GCCCTTGC	1626
OY	1612	GCGTGTCTCTCCACAGCTGACGTGACCTCACTGCGGTGTGTGATGATCACC GCCCTTGC	1671
Db	1627	TGCCCTGTATATGATATCGTATGTCAGAAACCTGTGTAGCTGACATCGGGGAAACAGATGA	1686
OY	1672	TGCCCTGTATATGATATCGTATGTCAGAAACCTGTGTAGCTGACATCGGGGAAACAGATGA	1731
Db	1687	AAGACAAAAAGAGAACCAAGTATATCTGCTTCCCTGCTTGTATATCTGAATAATCCAGTT	1746
OY	1732	AAGACAAAAAGAGAACCAAGTATATCTGCTTCCCTGCTTGTATATCTGAATAATCCAGTT	1791
Db	1747	TTAAAGCTCCGTTGAGAGACAGTTCACCCAATGSAACCTCTTCTGTTTAAAGAA	1806
OY	1792	TTAAAGCTCCGTTGAGAGACAGTTCACCCAATGSAACCTCTTCTGTTTAAAGAA	1851
Db	1807	AGCTGTGGCCCTCAGGGGGCTTGTGTGAAGAACTGCTACAGGGCTAATTTCCAAACCAT	1866
OY	1852	AGCTGTGGCCCTCAGGGGGCTTGTGTGAAGAACTGCTACAGGGCTAATTTCCAAACCAT	1911
	1867	AAGGCTTGGGGCGTGTGGGCTTAAGGGGACATTTGCACCATGTAAAGCAAGCTGGG	1926
	1912	AAGGCTTGGGGCGTGTGGGCTTAAGGGGACATTTGCACCATGTAAAGCAAGCTGGG	1971
Db	1927	CTTATCATGTGTGTGATGATGAGAGATGTAGTGTGATGATGATGTGTAATTTTAAACAGT	1986
OY	1972	CTTATCATGTGTGTGATGATGAGAGATGTAGTGTGATGATGATGTGTAATTTTAAACAGT	2031
Db	1987	TGAACCTGTCTCTCTACCTGTGTAGAGACAGAGATACCTATTGATAAAGATTTTCCAT	2046
OY	2032	TGAACCTGTCTCTCTACCTGTGTAGAGACAGAGATACCTATTGATAAAGATTTTCCAT	2091
Db	2047	GCTTACTACAGACAGANTGCTTCTGAAAGCAGGGCCGAGCC 2089	
OY	2092	GCTTACTACAGACAGANTGCTTCTGAAAGCAGGGCCGAGCC 2134	

RESULT	3	AF014117	2549 bp	mRNA	ROD	20-NOV-1997
LOCUS		Mus musculus	GDNF	receptor alpha (GDNFR-alpha)	mRNA, complete	cds
DEFINITION		AF014117				
ACCESSION		92624960				
NID						
KEYWORDS						
SOURCE		house mouse.				
ORGANISM		Mus musculus				
		Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
		Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

REFERENCE	1 (bases 1 to 2549)
AUTHORS	Dey,B.K., Mong,Y.W. and Too,H.P.
TITLE	Cloning of a novel murine isoform of the glial cell line-derived neurotrophic factor receptor
JOURNAL	Neuroreport 9 (1) (1998) In press
REFERENCE	2 (bases 1 to 2549)
AUTHORS	Dey,B.K., Mong,Y.W. and Too,H.P.
TITLE	Direct Submission
JOURNAL	Submitted (15-JUL-1997) Biochemistry, 10 Kent Ridge Crescent, Singapore 119760, Singapore
FEATURES	Location/Qualifiers

gene	1. .2549
	/gene="GDNFR-alpha"
CDS	914. .2320

/gene="GDNFR-alpha"  
/note="glial cell-line derived neurotrophic growth factor  
receptor"

translatio n "m"atly fuplyplldlmlmsvsgsdlldcvcakdmclokeosct /krtlarocvagrknfmsltgylekdkpckrsmeklksklyncrkcmkrekcklcti ymsyoslogndlieds pyepvsnrslndt rai pvtl sdyoo vhe lskgnlcladaa nca cindmckkyrsa yitpcttsmsnecvncnrpkckal rfofckpdkhysmlrscscd yacterrroqti ypcvcsyeerercnclnodsct nty ncrslad ffnocpessvsn clentadcllarlsglgiwtfnr yidsslvapmcdsgsdlldcclaflefrpd nltclknai oafngsdyttmopapvottatitlarl nklpdpagssneltlthv lpca nloaloklsvsgssthclcsdndygdgclagasslhttsmaap pscgslslpv mvtlaaolklsvslaets "

BASE COUNT	623 a	680 c	693 g	553 t
ORIGIN				

Query Match	Score	DB 23	length
75.5%	1615	2549	

Best Local Similarity 94.2%; Pred. No. 0.00e+00;  
Matches 1785; Conservative 0; Mismatches 98; Indels 12; Gaps 4;

D6 663 GGCAGTTGGGTCGGAACCTGAACCCCTAAAGCGGGTCCGCCCTCCGCCCTCGGGCCCGC 722

51 GGCAGTTGGTGGACTGAACCCCTGAAGCGGGTCCGCCCTCCGCCCTGGCGCCGC 110

Db 723 TCGAGCTGAGTCCCTGGCGCGGTGGCGGCGACAGCACGGGAGTCTGCTCTACACCT 782

111 CCGATCTGAGTCGCTGGCGGGGTGGGGCGCAGACGACGGGGAGTCTGCTCACCT 170

**Db 783 GGATGAGCTTAACCTTGTAGTGGCCAGAGGAGCGCAGTCGCCGGGGATCTGTGCACGCT 842**

171 GGATGAGCTGAACCTTTGACATGGCCAGAGGAGCGCAGTCGCCCGGGGATCGCTGCACGCT 230

Db 843 GAGCTCTCCGCGAGATCCGCTGCGGCTTTGATTTTGGGGGGGGGACAGCTGC 902

231 GAGCTCTCTCCCGAGACCGGGCGGGCGCTTTGATTTTGGGGGGGGCGGGGACAGCTGC 290

D6 903 GCGGGGACCATGTTCCTAGCCACTCTGTACTTGTGCTGCACCTCCTGATTTTGTA 962

291 GCGCGGCACCATGTTTCTAGCCACTCTGTACTTCGCGCTGCACCTCTTGATTTGCTGA 350

Db 963 TGTGGCCGAGTGTGTTGGGGACCGCTGACTGTGTGAAAGCCAGTATCAGTGC 1022

351 TCTCCGCCGAGGTGAGTGGTGGAGACCGTCTGAGCTGTGTGAAGCCAGCGATTCAGTGCC 410

Db 1023 TGAAGACAGAGCTGCAGCACCAAGTACCGCACACTGAGGCA6TGCGTGGCGGCAAGG 1082

411 TGAAGCAGACAGACTGCAGCACCAAGTACGGCACACTAAGGCAGTGCCTGGCGGGCAAGG 470

Db 1083 AAACCAACTTCAGCTGACATCCGGCTTCGAGGCCAAGATGAGTCCCGCAGCGCTATGG 1142

471 AAACCACTTCAGCCTGACATCCGGCCTTGAGGCCAAGATGATGCTGCCGTAGCGCCATGG 530

D	1143	AGGCTTGAAGCAGAACTCTCTCTACAACTGCCGCTGCAGAGGGGGGCAATGAAGAAAGCA	1202
O	531	AGGCTTGAAGCAGAACTCTCTCTCTACAACTGCCGCTGCAGAGGGGGGCAATGAAGAAAGCA	590
D	1203	AGAAATTTCTGCGATATCTACTGGAGCATGTACAGAGCCTGCAGGGAAATGACTACTAG	1262
O	591	AGAAATTTCTGCGATATCTACTGGAGCATGTACAGAGCCTGCAGGGAAATGACTACTAG	650
D	1263	AGAAATTTCCCAATACGAGCCGGTTAACAGCAGGCTGTCAATATATTCGGGGAGTCCCGT	1322
O	651	AGAAATTTCCCAATACGAGCCGGTTAACAGCAGGCTGTCAATATATTCGGGGAGTCCCGT	710
D	1323	TCATATTCAGATGTTTTCACCACTGGAAACACTTTTCCAAAGGAAACAACTGCCCTGATG	1382
O	711	TCATATTCAGATGTTTTCACCACTGGAAACACTTTTCCAAAGGAAACAACTGCCCTGAGC	770
D	1383	CAGCAAGAGCCTGCACCTGGATGACACCTGCAGAAAGTACAGTCCGCTCATCACC	1442
O	771	CAGCAAGAGCCTGCACCTGGAGCAGCACTGTAAAGATACAGTCCGCTCATCACC	830
D	1443	CCTGTACCAACAGCATGTCCAAATGAAGTGTCAACCGCCGCAAGTGCACAAAGCCTTCA	1502
O	831	CCTGTACCAACAGCATGTCCCAAGAGGTGTCAACCGCCGTAAGTGCACAAAGCCTTCA	890
D	1503	GGCAGTTCTTTCGACAAAGTTTCAGCCCAAGCAGCAAGCTACGAGATGCTCTCTCTCC	1562
O	891	GGCAGTTCTTTCGACAAAGTTTTCGCGCCCAAGCAGCAAGCTACGAGATGCTCTCTCTCC	950
D	1563	GGAGAGTCCGCTGTCACGAGAGGGGGGAGACAGACTATGCGCTGTGTCTCTCTAGAG	1622
O	951	GGAGAGTCCGCTGTCACCGAGGGGGGGGAGAGACTATGCTCCCTGTGTCTCTAGAG	1010
D	1623	AACGAGAGAGCCCACTGCTGTAATCTGCAGAGACTCTGCAGAGCAATTAATCATCTGCA	1682
O	1011	AACGAGAGAGCCCACTGCTGTAATCTGCAGAGACTCTGCAGAGCAATTAATCATCTGCA	1070
D	1683	GATCTCGCTTCGAGATTTTTTTTACCACTGCCAGCCAGATCAGGTCTGTACAGACT	1742
O	1071	GATCTCGCTTCGAGATTTTTTTTACCACTGCCAGCCAGATCAGGTCTGTACAGACT	1130
D	1743	GTCCTAAGGAACTATACGAGACGCTCCTGCGCTACGCGGAGACTGATGGCAGAGCA	1802
O	1131	GTCCTAAGGAACTATACGAGACGCTCCTGCGCTACTCGGAGCTGATGGCAGAGCA	1190
D	1803	TGACTCCCTAATACATAGACTCCAGCAGCCTCAGTGTGGCGCGTGTGTCAGTTCACACA	1862
O	1191	TGACTCCCAACTACTGTAGACTCCAGCAGCCTCAGGCTGGACAGATGTGTACTGTACACA	1250
D	1863	ACAGTGGCAATGACTGTGAGAAAGTTCCTGAAATTTTTTTTAAAGCAATACGT	1922
O	1251	ACAGCGGCAATGACTGTGAGAAAGCTGTGAAATTTTTTTTAAAGCAATACTT	1310
D	1923	GTCCTCAAAAATGCAATTCAGCCTTTTGGCAATGGCTCGGATGTGACCATGTGGCAACAG	1982
O	1311	GTCCTCAAAAATGCAATTCAGCCTTTTGGCAATGGCTCAATGTGACCATGTGGCAGCAG	1370
D	1983	CCCCCAGTCCAGACACCACTGTGCCAGCTACACTGCTTCGGATCAAGAAACAGC	2042
O	1371	CCCCCTCAGTCCAAACCACTGTGCCAGCTACACTGCTTCGGGTCAGAAACAGC	1430
D	2043	CTTAGGGCCAGCAGGCTCTGAGAAATGAGATTCACACACAGTTTATCCACCGTGTGTA	2102
O	1431	CTTAGGGGCCAGCAGGGCTGAGAAATGAGATTCACACACAGTTTATCCACCGTGTGGA	1490
D	2103	ATTTCGAGGCAACAGCTGAAATCCAAATATCCGGGAGTACACTCTCTCTTTCTG	2162
O	1491	ATTTCGAGGCTCAAACTGAAATCCCAATGTGTGCGGATGACACACACTCTCTCTTTCTG	1550
D	2163	ATAATGATTAGGAAAGGATGTGTGCTGTGTGCTGTCCAGCCAGCATTAACCAAAATCAA	2222
O	1551	ATAATGATTGTGGAAAGAGATGTGTGTGCTGTGTGTGTGTCCAGCCAGCATTAACCAAAATCAA	1610
D	2223	TGGCTGTCTCTCCAGCTGCGGCTGTGAAGTCACTGCGGCTGATGTTCACCGCTCTGG	2282

Qy	1611	TGGCTGCTCTCCACAGCTGACGTCTAGCTACTGCGGTGCTATAGCTACACGCCCTTG	1670
Db	2283	CTGCCCTGTTGTCTGTATCTATTCGCGAACAATCGTAGCTGCATCC-----AGTATG	2334
Qy	1671	CTGCCCTGTTATCTGTATCTGTTCTGTGGCAGAAACGTCGAGCTGCATCCGGGAAACAGATG	1730
Db	2335	AAAAGACAAAAGAAAACAAGTATTCTGTGCCCTGTCTCTCTGTATCTATAAATCCAG	2394
Qy	1731	AAAAGACAAAAGAAACCAAGTATTCTGTGCC-TGTCTCTGTATCTATAAATCCAG	1789
Db	2395	TTTTAAAAGCTCGGTTGAGACAGATTTACCCACGCGGAACTCCTTGTGTTTAAAG	2454
Qy	1790	TTTTAAAGCTCGGTTGAGACAGATTTACCCACTGSAACCTCTTCTGTTTAAAG	1849
Db	2455	AAAGCTTTTTCGCGCTCTACACAGCTTCTGTGTGAAGAACTGCTACAGGCTAAATCCAA	2514
Qy	1850	AAAGCTTGT-G--GCCCTACAGGGCTTCTGTGTGAAGAACTGCTACAGGCTAAATCCAA	1906
Db	2515	CTCAGAAAACCTGCGGCAAGTGTGTGGCTTCCAGGG	2549
Qy	1907	CCCAATAAGCTCTGGGGCTGTGGCGGCTTAAAGG	1941

[illegible]

Db	1	CCCCCTAAAAGCGGGTCCGGCTCCCGGCCCTTGCGGCCGCTCGAGCTGAATCCCTGGCGC	60
Oy	73	CCCCCTGAAAGCGGGTCCGGCTCCCGGCCCTTGCGGCCGCGGATCTGAATCCCTGGCGC	132
Db	61	GGTGGCGCGCAGACCAACGGGAGATCGCTCAACCCTGGATGAGACTTAATTGAGTG	120
Oy	133	GGTGGCGCGCAGACCAACGGGAGATCGCTCAACCCTGGATGAGACTTAATTGAGTG	192
Db	121	GCCAGAGAGCGCAGTCGCCCCGGGGATCTCTGACGCTGAGCTCTCTCCGCGAGATCCGG	180
Oy	193	GCCAGAGAGCGCAGTCGCCCCGGGGATCGCTGACGCTGAGCTCTCTCCCGAGACCGG	252
Db	181	TGGGGGCTTTGGATTTTGGGGGGGGGGGGGACAGCTGCGGGGGGACCATTTCTTAGC	240
Oy	253	CGGGGCTTTGGATTTTGGGGGGGGGGGGGACAGCTGCGGGGGGACCATTTCTTAGC	312
Db	241	CACCTGTACTTCTGCTGCGCACCTCTGTGATTTTGGCTGATGTGCGCGCAGAGTAGTG	300
Oy	313	CACCTGTACTTCTGCGCTGCGCACCTCTGTGATTTTGGCTGATGTGCGCGCAGAGT	372
Db	301	GGACCCGCTGGACCTGTGTGAAGCCAGTATATAGTGCCTTGAAGGAACAGACTGCAGAC	360
Oy	373	AGACCGCTGGACCTGTGTGAAGCCAGCAGTACGTCCCTGAAGGAACAGAGCTGCAGAC	432
Db	361	CAAGTACCGCACACTGAGGAGTGCCTGGCGGGGCAAGAAACCACTTGACCTGACATC	420
Oy	433	CAAGTACCGCACACTTAAGGCAGTGTGCTGGCGGGCAAGAAACCACTTGACCTGACATC	492
Db	421	CGGCTCGAGGCGCAAGATGATGTGCGGACGCTATGAGAGCCCTTGAAAGCAGAGTCT	480
Oy	493	CGGCTTGAGGCGCAAGAGTATGTGCGCTACGCGCAATGAGAGCCCTTGAAAGCAGAGTCT	552
Db	481	CTCAACACTGCCGCTGCAGAGGGGGCAATGAAGAAAGAAATTTGCTCTCGATATCTG	540
Oy	553	GTACAACTGCCGCTGCAGAGGGGGCATGAAGAAAGAAATTTGCTCTCGATATCTG	612
Db	541	GAGCATGTACAGAGCCTGCAGGGAAATGACCTACTGAGAAATTCGCCATACAGACCGGT	600
Oy	613	GAGCATGTACAGAGCCTGCAGGGAAATGACCTCTGAGAAATTCGCCATAGACCGGT	672
Db	601	TAAACGAGCGCTGCACATATATTTCCGGGAGTCCCGTTCAATATAG-----	647
Oy	673	TAAACGAGCGTTGCACATATATTTCCGGGAGTCCCGTTCAATATAGATTTTCCAGCA	732
Db	648	--TGGAACACATTTTCCAAAGGAAACAATCCCTCGATGCAAGCAAGGCTCGACCTGGA	705
Oy	733	AGTGAACACATTTTCCAAAGGAAACAATCGTGTGAGCGACGAAGGCTCGACCTGGA	792
Db	706	TGACACTCTGCAGAAAGTACAGTGTCCGCTTACATACCCCTGTACCAACAGCATTCGAA	765
Oy	793	CGACACCTGTAAAGAAATACAGTGTCCGCTTACATACCCCTGTACCAACAGCATTCGAA	852
Db	766	TGAAGTCTGCAACCGCGCAAGTGCACAAAGGCTCTAGGAGATTTTCTTGACAAAGTTC	825
Oy	853	CGAGTCTGCAACCGCGCTAAGTGCACAAAGGCTCTAGGAGATTTTCTTGACAAAGTTC	912
Db	826	AGCCAAGCAGACTAGGGGATGCTCTTGTGCTCTGTCCGGAGAGTGCCTGCACCGAGAG	885
Oy	913	GGCCAAGCAGACTAGCGGATGTGCTTCTCTCTCCGGAGATGCCTGCACCGAGAG	972
Db	886	GGGGGACAGACTATGTGCTGCTGTGTGCTCTTGAAGAAAGAGAGGCCCCAATGCT	945
Oy	973	GGGGGACAGACTATGTGCTGCTGTGTGCTCTTGAAGAAAGAGAGGCCCCAATGCT	1032
Db	946	GAATCTGCAGACACTCTGCAAGACAAATTACATCTGCAGATCTCGCTTGACGATTTTT	1005
Oy	1033	GAGTCTGCAAGACTCTGCAAGACAAATTACATCTGCAGATCTCGCTTGACGATTTTT	1092
Db	1006	TACCAACTGCCAGGCCAAGTCAAGGTCTGTGACGAACTGTCTTAAGGAGAACTACGCGA	1065
Oy	1093	TACCAACTGCCAGGCCAAGTCAAGGTCTGTGACGAACTGTCTTAAGGAGAACTACGCGA	1152

Db	1086	CTGCCTCTGGCCCTACTCTGGGACTGATTTGGCACAGTCATGACTCTCTACTACTATGACTC	11285
Oy	1153	CTGCCTCTCTGGCCCTACTCTGGGAGTGTATGGCACAGTCATGACTCTCTACTACTATGACTC	1212
Db	1126	CAGCAGCCTCAGTGTGGCGCCGATGGATGTCATGTCACACACAGTGGCAATGACTCTGAGAGA	11858
Oy	1213	CAGCAGCCTCAGCCTGGCCAGCACCATGATGTGTACTGTACAGCAACAGCGGCATGACTCTGAGAGA	12774
Db	1186	TTGGCTGGAAGTTTCTGAAATTTTTTAAAGCAGATACGTGTCTCAAAAATGCAATTCAAGC	1245
Oy	1273	CTGCTTGAAATTTCTGAAATTTTTTAAAGCAGATACGTGTCTCTCAAAAATGCAATTCAAGC	13333
Db	1246	CTTTGGCAATGGCTCGGATGTGACATGTGGGACCAACAGCCCCCAGTCACACACACAC	13005
Oy	1333	CTTTGGCAATGGCTCAGATGTGACATGTGGGACGAGCCCTTCACAGTCACACACACAC	13923
Db	1306	TGCCAGCATACACACTCTCTCCGGATCAAGAAACAGCCCTTAGGCGCAGCAGGCTCTGA	1365
Oy	1333	TGCCACCATCTACACTCTCTCTCCGGGTCAGAAACAGCCCTCTGGGGCCAGCAGGCTCTGA	1453
Db	1366	GAATGAGATTTCCACACACGTTTTTACACCGTGTGCTAATTTGCAGGCAACAGACTGAA	14235
Oy	1453	GAATGAGATTTCCACACACGTTTTTACACCGTGTGGAATTTGCAGGCTCAAGAACTGAA	1512
Db	1426	ATCCAAATATACGGGCGAGTACACATCTCTGTCTTTCTGTAATGATATAGGAAAGATAGG	14885
Oy	1513	ATCCAAATGTGTGGGATGAGCACACACCTCTGTCTTTCTGTAATGATATTTGGAAAGATAGG	15727
Db	1486	TCTGCGTGGTACCTCCAGACCAATTAACCAAAATCAATGGCTGCTCCCTCCAGCTGGG	15454
Oy	1573	TCTGCGTGTGTGCTCCAGCCCAATTAACCAAAATCAATGGCTGCTCCCTCCAGCTGGAG	16333
Db	1546	TCTGAGCTCACTGGCGGTGATGGTGTTCACACCGCTCTGGCTGGCTGTGTCTGATTCAT	16005
Oy	1633	TCTGAGCTCACTGGCGGTGATGGTGTTCACACCGCTCTGGCTGGCTGTGTCTGATTCAT	16927
Db	1606	GGCAGAAACATCTTACGTGCATCC	1629
Oy	1693	GGCAGAAACGTCTGAGCTGCATCC	1716

```

RESULT      5      24-OCT-1997
LOCUS       RN2072      1392 bp      RNA
DEFINITION  Rattus norvegicus mRNA for GDNFR-alpha/Tnnr1-delta protein,
ACCESSION   AJ002072
NID         G2564301
KEYWORDS    GDNFR-alpha/Tnnr1-delta protein.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
             Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1      (bases 1 to 1392)
AUTHORS     Zhong,J.
TITLE       Direct Submission
JOURNAL     Submitted (16-OCT-1997) Zhong J., Molecular Neurobiochemistry
            NCJ/117, Ruhr University Bochum, Universitaetsstr. 150, D-44780
            Bochum, GERMANY
REFERENCE   2      (bases 1 to 1392)
AUTHORS     Zhong,J., Anlies,M. and Heumann,R.
TITLE       Rattus norvegicus GDNFR-alpha/Tnnr1-delta mRNA, complete cds
JOURNAL     Unpublished
FEATURES    Location/Qualifiers
             1..1392
             /organism="Rattus norvegicus"
             /strain="Hannover"
             /db_xref="taxon:10116"
             1..1392
             /codon_start=1
             /product="GDNFR-alpha/Tnnr1-delta protein"
             /db_xref="PID:e1154274"
             /db_xref="PID:g2564302"
             /translation="MPLATLVFALPLDLMSAEVSGDRIDCVKASDQCLEQSCST

```

KYRTLRQCVAGKETNFSLTSGLEAKDECRSMEALKOKSLYNCRKRGMKKELRLI  
 YMSYOSILOGNDLEDSPYENRSLRDIAPRAVEFISEVHISKNRNLDAKACINDD  
 TCKRYRATITPCTPSTMSNEVNRKRKHLAROFEDVPAKHSHGMLFCRCRODLACTE  
 RRROTIVPCSYIERERPNCLSLDSCRTINTICRSLADPFTNCQPSRVSNCLEEN  
 YADOLLAISGLIGVTMPNPNYDSSLSVAPMCDSCNSGNDLEDLKLNFEDKDTCLK  
 NATOAFNGSDVYTMQAPPYQVOTATTAFTAFKRNKPLGANGSENEIPIHVPPCAN  
 LOGOKLSNVSSTHLCISDSDFGKDLGASSHITTKSMAPSPCSLSLPLMLTA  
 LAALISVLSLETSS"

BASE COUNT 338 a 406 c 356 g 292 t  
 ORIGIN  
 Query Match 60.9%; Score 1302; DB 23; Length 1392;  
 Best Local Similarity 98.9%; Pred. No. 0.00e+00;  
 Matches 1392; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

Db 1 ATGTTCTAGCACTCTGTACTTGGCGTCCACTCTGTGATTTGCTGATCGCCGAG 60  
 Oy 302 ATGTTCTAGCACTCTGTACTTGGCGTCCACTCTGTGATTTGCTGATCGCCGAG 361  
 Db 61 GTGAGTGTGAGAGCCGTCTGTACTTGTGAAAGCCAGCATGCTGTGAAGAAAG 120  
 Oy 362 GTGAGTGTGAGAGCCGTCTGTACTTGTGAAAGCCAGCATGCTGTGAAGAAAG 421  
 Db 121 AGCTGACACCAAGTATACGACACTAAGGAGTCTGTGCGGCGCAAGAAACCACTTC 180  
 Oy 422 AGCTGACACCAAGTATACGACACTAAGGAGTCTGTGCGGCGCAAGAAACCACTTC 481  
 Db 181 AGCTGACATCCGCGCTTGAGGCCAAGATGATGTCGATGCGCATGAGGCGCTTGAAG 240  
 Oy 482 AGCTGACATCCGCGCTTGAGGCCAAGATGATGTCGATGCGCATGAGGCGCTTGAAG 541  
 Db 241 CAGAAGTCTGTACTGCTGCGGTGCAAGGCGGCGATGAAGAAAGAAAGAAATGTGTG 300  
 Oy 542 CAGAAGTCTGTACTGCTGCGGTGCAAGGCGGCGATGAAGAAAGAAAGAAATGTGTG 601  
 Db 301 CGTATCTACTGAGAGCATATACAGAGCCTGCAAGGAAATGACCTCTGGAAGATTCGCCG 360  
 Oy 602 CGTATCTACTGAGAGCATATACAGAGCCTGCAAGGAAATGACCTCTGGAAGATTCGCCG 661  
 Db 361 TATGAGCCGGTTAACAGAGTGTGCAATATATTCGCGGCGAGTCCGCTCATATACAG-- 418  
 Oy 662 TATGAGCCGGTTAACAGAGTGTGCAATATATTCGCGGCGAGTCCGCTCATATACAGAT 721  
 Db 419 -----TGACACACTTCCAAAGGAAAGCAACTGCTGTGAGAGCCCAAGGCC 465  
 Oy 722 GTTTCAGCAAGTGTGAACACTTTCCAAAGGAAAGCAACTGCTGTGAGAGCCCAAGGCC 781  
 Db 466 TGCAACTGAGAGACACCTGTGAAGAGTACAGGTGCGGCTCATACCCCGTCGACACAC 525  
 Oy 782 TGCAACTGAGAGACACCTGTGAAGAGTACAGGTGCGGCTCATACCCCGTCGACACAC 841  
 Db 526 AGCATGTCCAAGAGGTGTGCAACCGCGTGAAGTGCACAAAGGCCCTCAGGCAATGTTCTC 585  
 Oy 842 AGCATGTCCAAGAGGTGTGCAACCGCGTGAAGTGCACAAAGGCCCTCAGGCAATGTTCTC 901  
 Db 586 GACAAAGTTCGCGGCGCAAGACAGTACGGGATGCTCTCTGCTCGCGGAGCAATCGCC 645  
 Oy 902 GACAAAGTTCGCGGCGCAAGACAGTACGGGATGCTCTCTGCTCGCGGAGCAATCGCC 961  
 Db 646 TGCACCGAGCGCGCGAGACAGTATCGCCCGTGTCTCTATGAAGAAAGAGAGAG 705  
 Oy 962 TGCACCGAGCGCGCGAGACAGTATCGCCCGTGTCTCTATGAAGAAAGAGAGAG 1021  
 Db 706 CCAACTGCTGAGTCTGCAAGACTCTGCAAGACCAATTCATCTGAGATCTGCGCTT 765  
 Oy 1022 CCAACTGCTGAGTCTGCAAGACTCTGCAAGACCAATTCATCTGAGATCTGCGCTT 1081  
 Db 766 GCAGATTTTTCACCAATTCGCAAGAGTCAAGTCTGTAGCAACTGTCTTAAGAG 825  
 Oy 1082 GCAGATTTTTCACCAATTCGCAAGAGTCAAGTCTGTAGCAACTGTCTTAAGAG 1141  
 Db 826 AACATGAGAGTGTGCTGCTGCTACTGCGGAGTGTGAGAGTGCATGATGCCAAC 885

Oy 1142 AACTAGCAGACTGCTCTCTGCGCTACTCGGAGATGATGGACAGTCAATGCCAAC 1201  
 Db 886 TACCTAGACTCCACACAGCTTACGAGTGTGGACACATGAGTGTACTGACACACAGGGCAAT 945  
 Oy 1202 TACCTAGACTCCACACAGCTTACGAGTGTGGACACATGAGTGTACTGACACACAGGGCAAT 1261  
 Db 946 GACCTGGAAGACTGCTGAAATTTCTGAAATTTTAAAGCAATFACCTGTCTCAAAAAT 1005  
 Oy 1262 GACCTGGAAGACTGCTGAAATTTCTGAAATTTTAAAGCAATFACCTGTCTCAAAAAT 1321  
 Db 1006 GCAATTCAGCCCTTGGCAATGCTCAGATGTGACATGTGGACAGCCGCTCCAGTC 1065  
 Oy 1322 GCAATTCAGCCCTTGGCAATGCTCAGATGTGACATGTGGACAGCCGCTCCAGTC 1381  
 Db 1066 CAGACACACACTGCGCACACCTACACATGCTCTCGGGGTCAAGAAACCTCTGGGCGCA 1125  
 Oy 1382 CAGACACACACTGCGCACACCTACACATGCTCTCGGGGTCAAGAAACCTCTGGGCGCA 1441  
 Db 1126 GCAGGGTCTGAGATGAGATCCCAACAGCTTTTACACACCTGTGGGAATTTGCAGGCT 1185  
 Oy 1442 GCAGGGTCTGAGATGAGATCCCAACAGCTTTTACACACCTGTGGGAATTTGCAGGCT 1501  
 Db 1186 CAGAAGTGAATTCGAATGTGTGCGGTAGACACACACCTGTCTTTCTGTGATGATGATTC 1245  
 Oy 1502 CAGAAGTGAATTCGAATGTGTGCGGTAGACACACACCTGTCTTTCTGTGATGATGATTC 1561  
 Db 1246 GGAAGAGTGTCTGCTGCTGCTGCTCCAGCCACATTAACCAAAATCAATGCTGCTCT 1305  
 Oy 1562 GGAAGAGTGTCTGCTGCTGCTGCTCCAGCCACATTAACCAAAATCAATGCTGCTCT 1621  
 Db 1306 CCGAGCTGCACTGAGCTCACTGCGGTGCTGATGCTCAACCGCCCTTGTGCTGCTGCTTA 1365  
 Oy 1622 CCGAGCTGCACTGAGCTCACTGCGGTGCTGATGCTCAACCGCCCTTGTGCTGCTGCTTA 1681  
 Db 1366 TCTGTATCTGTGGCAGAAACGTCGTAG 1392  
 Oy 1682 TCTGTATCTGTGGCAGAAACGTCGTAG 1708

RESULT 6  
 LOCUS AB000800 1415 bp mRNA ROD 04-FEB-1997  
 DEFINITION Mouse mRNA for GDNF receptor alpha, complete cds.  
 ACCESSION AB000800  
 NID g1816441  
 KEYWORDS GDNF receptor alpha.  
 SOURCE Mus musculus Dorsal root ganglion Schwann cell cell\_line:IMS32 cDNA to mRNA.  
 ORGANISM Mus musculus  
 Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1415)  
 AUTHORS Watabe,K.  
 TITLE Regulation of GDNF receptor alpha mRNA in Schwann cell cultures  
 JOURNAL Unpublished (1997)  
 REFERENCE 2 (bases 1 to 1415)  
 AUTHORS Watabe,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JAN-1997) to the DDBJ/EMBL/GenBank databases.  
 Kazuhiko Watabe, Tokyo Metropolitan Institute for Neuroscience, Neuropathology, Musashidai 2-6, Fuchu, Tokyo 183, Japan  
 (E-mail:kazut@metn.ac.jp, Tel:0423-25-3881, Fax:0423-21-8678)  
 FEATURES  
 Location/Qualifiers  
 1. 1415  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_line="IMS32"  
 /cell\_type="Schwann cell"  
 /tissue\_type="Dorsal root ganglion"  
 1. 20  
 2. 1408  
 primer\_bind  
 CDS  
 /standard\_name="glial cell line-derived neurotrophic factor receptor alpha"









||||| 1846 TAAGAAAGCTTGT-G-GC-CCTCAGGGGCTTCTTTGAAGAACTGACAGGGGCTATTTC 1902  
Db 2160 CAAACTCAGAAAGCTTTGGGATATGCTGATTTTAAAGGACACTTTGACTTG 2214  
QY 1903 CAAACCCATTAAGGCTCTGGGGCGGTGGCTTAAGGGACATTTGCACCATG 1957

LOCUS 8 AF042080 1619 bp mRNA PRI 22-JAN-1998  
DEFINITION Homo sapiens glial cell line-derived neurotrophic factor receptor  
ACCESSION AF042080  
NID 92801556  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE  
1 (bases 1 to 1619)  
Primates; Catarrhini; Hominoidea; Homo.  
2 (bases 1 to 1619)  
Shenolik, S.E., Khorana, S., Schultz, P.N., Huang, E., Thobe, N.,  
Hu, Z.J., Fox, G.M., Jing, S., Cole, G.J., and Gagel, R.F.  
Mutational analysis of the GDNF/RET-GDNF signaling complex in a  
kindred with vesicoureteral reflux  
Hum. Genet. (1998) in press  
3 (bases 1 to 1619)  
Shenolik, S.E., Khorana, S., Schultz, P.N., Huang, E., Thobe, N.,  
Hu, Z.J., Fox, G.M., Jing, S., Cole, G.J., and Gagel, R.F.  
Direct Submission  
Submitted (08-JAN-1998) Endocrinology-Box 15, M.D. Anderson Cancer  
Center, 1515 Holcombe Blvd, Houston, TX 77030, USA

FEATURES  
source  
1. 1619  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q25-q26"  
/tissue="type=medullary thyroid carcinoma tumor"  
1. 1619  
/note="GDNFRa"  
/gene="GFRa1"  
/note="GFRa1"  
/note="GFRa1-linked receptor"  
/product="glial cell line-derived neurotrophic factor  
receptor alpha"  
/db\_xref="pid:g2801557"  
/translation="MFLATLYFALPLDLILSAEVSQDRLDCVKASDQCLKEQSCST  
KRTTLCVAGKTESLASGLAEKDECSAMALOKSLVRCRGRKREKNCARI  
VMSVOSLOGNDLEDSPEPNSRLSDIFRVPFISDVOVEHPRKGNLDAKA  
CHUDICKRYSRITPTCTSVSNDCNRKCKALROFDVPAHSHSGMLFCSDRD  
LCTERKROTIVPVCSEIEREKPNCLNDDSCNTNCSRLADEFTNCQPSRVS  
CKENKADCLAFSGISGVTMPTNIDSLVAPWDCNSNGNDEDLCKPESRVS  
NCLKNAIOAFGNSDVTWOPAFVOTATATTTATLARKNPLGAGSENIPTVL  
PCANAOAKLKSNSVNSGTHLCLISNRYKEGIGASHTTSMAPPCGSLPLVL  
VYATLSTLSTETS"  
529..543  
misc\_feature  
/gene="GFRa1"  
/note="found to be alternative exon after partial analysis  
of genomic sequence"

BASE COUNT 409 a 448 c 419 g 343 t  
ORIGIN

Query Match 53.3%; Score 1139; DB 22; Length 1619;  
Best Local Similarity 86.5%; Pred. No. 0.00e+00;  
Matches 1404; Conservative 0; Mismatches 211; Indels 9; Gaps 5;

Db 1 CAAAGCAGCAGCAGCTGTCGGGGGATGCTGCAGGCTAGCTCCCTCGGCAAGACCCAGCG 60  
QY 195 CAAAGAGCCAGCTGCGCGGGGATGCTGCAGGCTAGCTCCCTCGGCAAGACCCAGCG 254  
Db 61 GCGGCTGCGGATTTTGGGGGGGGGAGACAGCCCGCGCGGACCATGTTCTG 120

QY 255 GCGGCTTGGATTTT---GGGGGGGGGGGACAGCTGCGGGGGGACCATTTCTTAG 311  
Db 121 CGACCTGTACTTGGCTGCGCTGCTTGGACTTCTCTGTCGCGGAGTAGCGGGC 180  
QY 312 CCACTGTACTTGGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 371  
Db 181 GAGACCGCTGAGTGTGTAAGGACGATGATGATGATGATGATGATGATGATGATG 240  
QY 372 GAGACCGCTGAGTGTGTAAGGACGATGATGATGATGATGATGATGATGATGATG 431  
Db 241 CCAAGTACCGCAGCTTAAGGACGATGATGATGATGATGATGATGATGATGATGATG 300  
QY 432 CCAAGTACCGCAGCTTAAGGACGATGATGATGATGATGATGATGATGATGATGATG 491  
Db 301 CCGGCTGAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
QY 492 CCGGCTGAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 551  
Db 361 TCTAACACTGCGCTGAGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 552 TCTAACACTGCGCTGAGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 611  
Db 421 GGAGCATGTACAGAGCTGCAAGGAAATGATGATGATGATGATGATGATGATGATGATG 480  
QY 612 GGAGCATGTACAGAGCTGCAAGGAAATGATGATGATGATGATGATGATGATGATGATG 671  
Db 481 TTAACAGCAGATGTGATATATTCGGGGTGTCCCATTCATATGATGATGATGATGAT 540  
QY 672 TTAACAGCAGATGTGATATATTCGGGGTGTCCCATTCATATGATGATGATGATGAT 721  
Db 541 AAGTGAAGCATTCCCAAGGAAACATGCTGATGATGATGATGATGATGATGATGATGATG 600  
QY 732 AAGTGAAGCATTCCCAAGGAAACATGCTGATGATGATGATGATGATGATGATGATGATG 791  
Db 601 ACAGCATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
QY 792 ACAGCATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 851  
Db 661 ACAGCATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
QY 852 ACAGCATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 911  
Db 721 CGGCAAGCAGCAGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
QY 912 CGGCAAGCAGCAGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971  
Db 781 GGAGGAGCAGCAGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
QY 972 GGAGGAGCAGCAGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1031  
Db 841 TGAATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 1032 TGAATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1091  
Db 901 TTAACAGCAGCAGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
QY 1092 TTAACAGCAGCAGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1151  
Db 961 ACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1152 ACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211  
Db 1021 CCAAGTACCGCAGCTTAAGGACGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
QY 1212 CCAAGTACCGCAGCTTAAGGACGATGATGATGATGATGATGATGATGATGATGATGATG 1271  
Db 1081 ACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
QY 1272 ACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331  
Db 1141 CTTTGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

1332 CCTTGGCAATGGCTCAGATGTGACCATGTGGCAGCCGCCCTCCAGTCCAGACCACCA 1391  
Db 1201 CTGGCACTACCAACCATGGCCCTCCGGGTTAAGAAACAGCCCTGGGGCCAGCGGCTGG 1260  
1392 GTGGCAACCACTACCTGCTCCGGGTCAAGAAACAGCCCTGGGGCCAGCGGCTGG 1451  
Db 1261 AGAATGAATTCCTCACTATGTTTTCACACCGGTGCAAAATTTACAGGACAGAACTGA 1320  
1452 AGAATGAATTCCTCACTATGTTTTCACACCGGTGCAAAATTTACAGGACAGAACTGA 1511  
Db 1321 AATCAATGTGTGCGGCAATACACACCTGTATTTCCAAATGATTAATGAAAAAAG 1380  
1512 AATCAATGTGTGCGGCAATACACACCTGTATTTCCAAATGATTAATGAAAAAAG 1571  
Db 1381 GTCTCG---GTCTTCACGACCAATACCAAAATCAATGGTGTCTCTCCAAAGTGTG 1437  
1572 GTCTCGCTGGTGTCTCCACGACCAATACCAAAATCAATGGTGTCTCTCCAAAGTGTG 1631  
Db 1438 GTCTGAGCCCACTGTGTCTCTCTGTAACGCTGTCTGTCACCAATTAATCTTTAAAG 1497  
1632 GTCTGAGCCCACTGTGTCTCTCTGTAACGCTGTCTGTCACCAATTAATCTTTAAAG 1691  
1498 AATCATCATGCTGATTAATAAATAATATGACATGTAAAAAGACAA- AACCAAG 1556  
1692 TGCGAAGAAAGCTGTGATGATGATCGGGAACAGATGAAAAAGACAAAGAACCAAG 1751  
Db 1557 TATCTGTTTCTGTTCTCTCTGTAATAGCTGAATTTCCAGTTT- AGAAGCTCACTTGAAG 1615  
1752 TATCTGTTTCTGTTCTCTCTGTAATAGCTGAATTTCCAGTTT- AGAAGCTCACTTGAAG 1810  
Db 1616 ACAG 1619  
1811  
1811 GCAG 1814

RESULT 9 HSN95847 2175 bp mRNA PRI 02-OCT-1997  
LOCUS Human GDNF receptor alpha mRNA, complete cds.  
DEFINITION U95847  
ACCESSION 92459741  
NID  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2175)  
AUTHORS Hishiki,T., Kondo,K., Ichimiya,S., Nimura,Y., Seki,N., Ozaki,T.,  
Sakiyama,S., Takahashi,H., Ohnuma,N., Tanabe,M., Fujimura,S. and  
Nakagawa,A.  
TITLE GDNF-induced differentiation and its enhancement by retinoic acid  
in primary human neuroblastomas expressing c-Ret and GDNF-alpha  
Unpublished  
2 (bases 1 to 2175)  
JOURNAL Nakagawa,A.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (31-MAR-1997) Division of Biochemistry, Chiba Cancer  
Center Research Institute, 666-2, Nitona, Chuoh-ku, Chiba 260,  
Japan

FEATURES  
source  
1. 2175  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/dev\_stage="adult"  
/tissue\_type="substantia nigra"  
245..1627  
/codon\_start=1  
/product="GDNF receptor alpha"  
/db\_xref="PID:92459743"  
/translation="MFLATLYRALPLLDLLLSAEVSGGDRDCVAKSDQCKECSCT  
KYRLQCAVKEINFLASGLSEARDECSAMEALQKSLYKRCRKGKKEKNCRLI  
TWSMYVLOGNDLDEPYNRSLSDFRVPVPISEVHIGKNGKNCDAACMLDI  
ICKYRSATIPCTTSVNDVCRNRCHALDFDKYPAKSYMLPCSDIACTE  
RRROTIVPCSYSEERKPCNCLDLSCKRNYICRRLADFLFNCQPESSVSSCLKEN

BASE COUNT 544 a 566 c 535 g 516 t 14 others  
ORIGIN  
YADCLLAATSLIGTVMTPNYIDSSSLVAAPKDCDSNSGNDLECLKFNFKDNTCK  
NAIAPFGNSDVTWQAPFQVOTTTATTTATLRKNKPLGPAAGSENEIPTHLPKPN  
LQAKKSNVSGNTHLCLSNQNYEKELGASHTTTSMAPPCGLPLLVVLTAL  
STLSLDEITS  
Query Match 52.8%; Score 1128; DB 22; Length 2175;  
Best Local Similarity 83.0%; Pred. No. 0.00e+00;  
Matches 1634; Conservative 0; Mismatches 300; Indels 35; Gaps 17;  
Db 1 TGGGTGAGCACTGAACCCCTTAAGAGCGACCCCTCCCGCTCGC--CATCCGAGAC 58  
117 TGGGTGAGCACTGAACCCCTTAAGAGCGACCCCTCCCGCTCGC--CATCCGAGAC 117  
58 TGGGTGAGCACTGAACCCCTTAAGAGCGACCCCTCCCGCTCGC--CATCCGAGAC 117  
Db 59 TGAATGCGCGCGCGCGGTGG--CTGCTCCAGACCCGGAATTTCTTCTTCACTGATGA 117  
118 TGAATGCGCGCGCGCGGTGG--CTGCTCCAGACCCGGAATTTCTTCTTCACTGATGA 117  
118 GCTGAATTTGGGCGCGCGAGACACACACGCTGTCGGGGATGCTGATGCTGATGCTC 177  
178 GCTGAATTTGGGCGCGCGAGAGAGCGAGTCCCGGGGATGCTGATGCTGATGCTC 237  
Db 178 CTGCGCAAGACCCAGACCGCGCTCGGATTTTGGGGGGGCGGAGACACCCCGCGC 237  
238 CTGCGCAAGACCCAGACCGCGCTCGGATTTTGGGGGGGCGGAGACACCCCGCGC 294  
238 CTGCGCAAGACCCAGACCGCGCTCGGATTTTGGGGGGGCGGAGACACCCCGCGC 294  
238 CGGACACATGCTCTGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 297  
295 CGGACACATGCTCTGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354  
298 GCGCAAGTGAAGCGCGCGAGACCGCGCTGATGCTGTAAGCGCATGATGCTGCTGA 357  
355 GCGCAAGTGAAGCGCGCGAGACCGCGCTGATGCTGTAAGCGCATGATGCTGCTGA 414  
Db 358 GGAGCAGACCTGACACACCAATGACCGACGCTAAGGCGCTGCGGGGCGGAGAGAC 417  
415 GGAGCAGACCTGACACACCAATGACCGACGCTAAGGCGCTGCGGGGCGGAGAGAC 474  
418 CAATCTGACCTGCGATCCGCGCTGAGGCGCAAGATGATGCTGCGGCGCATGAGAGC 477  
475 CAATCTGACCTGCGATCCGCGCTGAGGCGCAAGATGATGCTGCGGCGCATGAGAGC 534  
478 CCGTAAGCAGAGTGGCTTCACTGCGCGCTGCAAGCGGGGTATGAAGAAGAGAA 537  
535 CTTGAAGCAGAGTGGCTTCACTGCGCGCTGCAAGCGGGGTATGAAGAAGAGAA 594  
538 CTGCGTGGCATTCTGAGCATGTACAGAGCTGCGAGAAATGATCTGCTGAGAGA 597  
595 TGTCTGCTATCTACTGAGCATGTACAGAGCTGCGAGAAATGATCTGCTGAGAGA 654  
598 TTCCCATATGACAGCATTTACAGAGATTTGCAATATATTCGGGTGCTCCATTCAT 657  
655 TTCCCATATGACAGCATTTACAGAGATTTGCAATATATTCGGGTGCTCCATTCAT 714  
658 ATCAG-TG-----G--AG-----CAATTTCCCAAGGAACTGCTGCTGAGAGC 702  
715 ATCAGATGTTTTCAGAGCATTTACAGAGATTTGCAAGGAACTGCTGAGAGAGC 774  
703 GAAGGCTGCAACCTGAGACATTTGCAAGAGTACAGAGTGGGCTGATCACCCCGTG 762  
775 CAAGGCTGCAACCTGAGACATTTGCAAGAGTACAGAGTGGGCTGATCACCCCTTG 834  
Db 763 CACCAACACAGCTGTCCAAATGATGTGCAACCGCGCAAGTGCACAAAGCCCTCGGCA 822  
835 CACCAACACAGCTGTCCAAATGATGTGCAACCGCGCTAAGTCCCAAGGCGCTCGGCA 894  
823 GTTCTTTGCAAGGTCCGGGCAAGCAGCATGCTGCTTCTGCTGCTGCTGCTGCTG 882  
895 GTTCTTTGCAAGGTCCGGGCAAGCAGCATGCTGCTTCTGCTGCTGCTGCTGCTG 954  
Db 883 CATGCGCTCAGAGAGCGGAGCGAGACCATGCTGCTGCTGCTGCTGCTGCTGCTG 942

```
Oy 955 CATGCGCTGCACCGGCGGCGGCGACAGACTATGCTCCCGGTGCTCTCTATGAAAGAC 1014
Db 943 GGAGAGCCCACTCTTTGAATTTTCAGAGACTCTGCAAGAAATTACATCTGCAGATC 1002
Oy 1015 AGAGAGGCCCACTGCTGAGCTGCAAGACTCTGCAAGAAATTACATCTGCAGATC 1074
Db 1003 TCGCGTTCGAGATTTTATACCACTGCGAGCGAGAGCAAGGCTGTCAGAGAGTCT 1062
Oy 1075 TCGCTTTCAGATTTTATACCACTGCGAGCGAGAGCAAGGCTGTCAGAGAGTCT 1134
Db 1063 AAAGGAAACTACGCTGACTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122
Oy 1135 TAAGAGAACTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1194
Db 1123 CCCCACTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1182
Oy 1195 TCCCACTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1254
Db 1183 TGGGAAAGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1242
Oy 1255 CGGCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1314
Db 1243 TAAAAATGCAATTCAGGCTTTGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1302
Oy 1315 CAAAAATGCAATTCAGGCTTTGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
Db 1303 CCCAATACAGAGACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1362
Oy 1375 TCCCACTACAGAGACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1434
Db 1363 GGGGCGACAGAGGCTGAGAAATGAAATTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422
Oy 1435 GGGGCGACAGAGGCTGAGAAATGAAATTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494
Db 1423 ACAGGACAGAGAGTGAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482
Oy 1495 GCAGGCTCAGAGAGTGAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1554
Db 1483 TAATTAATGAAAGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1539
Oy 1555 TGATTTGGAAGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1614
Db 1540 TGCCTCTCAAGAGTCTGCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599
Oy 1615 TGCTCTCTCCAGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 1674
Db 1600 CCTATTATCTTTACAGAAACATCATAGCTGATTAAAAAATACATATGAGCATGTAA 1659
Oy 1675 CCTATTATCTTTACAGAAACATCATAGCTGATTAAAAAATACATATGAGCATGTAA 1734
Db 1660 AAAGCAAAA-AMCAAGTATCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1717
Oy 1735 GACAAAAGAAACCAAGTATCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1793
Db 1718 AGAGCTAGTGTGAGAAACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1777
Oy 1794 AAAGCTCTGTTGAGAAACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1849
Db 1778 AAAGCTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCT 1837
Oy 1850 AAAGCTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCT 1906
Db 1838 CTCAAAAGGCTTGGGAGATGCTGTTGTTTAAAGGACAGTGTGTTGTTGTTGTTGTTGTTGTT 1897
Oy 1907 CCAATTAAGGCTTGGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1966
Db 1898 NCAATCTGGGAGTGTGTTTTC-GATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1945
Oy 1967 CTGGGCTATCATGTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2015
```

```
RESULT 10 HSU97144 1707 bp mRNA PRI 29-JUL-1997
LOCUS
```

```
DEFINITION Homo sapiens RET ligand 1 (RET1) mRNA, complete cds.
ACCESSION U97144
NID 92282025
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Carnivora; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 1707)
AUTHORS Sanicola,M., Hession,C.A., Morley,D.S., Carmillo,P., Ehrenfels,C.,
Malus,L., Robinson,S., Jaworski,G., Wei,H., Tizard,R., Whitley,A.,
Pepinsky,R.B. and Cate,R.L.
Glial cell line-derived neurotrophic factor-dependent RET
activation can be mediated by two different cell-surface accessory
proteins
Proc. Natl. Acad. Sci. U.S.A. 94 (12), 6238-6243 (1997)
JOURNAL MEDLINE 97322356
REFERENCE 2 (bases 1 to 1707)
AUTHORS Sanicola,M., Hession,C.A., Morley,D.S., Carmillo,P., Ehrenfels,C.,
Malus,L., Robinson,S., Jaworski,G., Wei,H., Tizard,R., Whitley,A.,
Pepinsky,R.B. and Cate,R.L.
Direct Submission
Submitted (11-APR-1997) Molecular Genetics, BIOGEN, 14 Cambridge
Center, Cambridge, MA 02142, USA
FEATURES
source
1..1707
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="embryonic"
/tissue_type="kidney"
1..1707
/gene="RET1"
262..1644
/gene="RET1"
/feature="PI-linked cell-surface accessory protein; GDNFR"
/codon_start=1
/db_xref="PI-linked cell-surface accessory protein; GDNFR"
/db_xref="PI-linked cell-surface accessory protein; GDNFR"
/translation="MFLATITFALPLDLILSLAEVGGDRDLCVKAASDCLAKROSGT
KYRFLQCVAGKSTNLSLAEVGGDRDLCVKAASDCLAKROSGT
WMSYOSLOGNEDLSDPYEPVNSRSLDFRVPFLVSEIPIRNCCLPAKACNLDI
ICKRYSAITPTCTVSNDVNCNRKCHKLRFEPVPAKHSYGLFCGRDIACIE
RRROTIVPVCYEREPKNCNLQDSCKTNYIRFSLRDLADFTNCOPESSRSCLKN
YACILAYSLGICTVMTPNYIDSSISLVAPMDCSSNGLERCLFLNFRKNTCLK
LNAOFKNGSDVTVKOPAPFVOTTTATTTTALVKKPKLPAPSEDEIPHYVPCAN
LNAOKLKSNSVGTHTLCTISNGNKEKGLASSHTTKSMAPSPGLPLVLVITAL
STLSLSTETS"
BASE COUNT 420 a 490 c 456 g 341 t
ORIGIN
Query Match 51.1%; Score 1093; DB 22; Length 1707;
Best Local Similarity 87.2%; Pred. No. 0.00e+00;
Matches 1427; Conservative 0; Mismatches 184; Indels 25; Gaps 8;
Db 15 GTTGGTGGAGCTGGAACCCCTAAAGGGAAGCGCTCCCGCTGCGC-CATCCCGGA 72
Oy 56 GTTGGTGGAGCTGGAACCCCTGAAAGGAGGTCCGCTCCCGCTGCGCGCCCGGGA 115
Db 73 GCTGAGTGGCGGGGGGGGTGG-CTGCTGCGAGACCGGAGTTTCTCTTCACTGATG 131
Oy 116 TGTGAGTGGTGGCGGGGGGTGGCGAGAGCGAGGGAAGTGTGCTCACTGATG 175
Db 132 GAGCTGAACCTTGGGGCGCCAGAACAGACAGTGTCCGGGAGTGTGCTGACGCTGAGCT 191
Oy 176 GAGCTGAACCTTGGAGTGGCCAGAGAGCGAGTCCCGGGGATCGCTGACGCTGAGCT 235
Db 192 CCGTGGGAAGACCGAGCGGCGCTCGGAGTTTGTGGGGGGGGGGGACGACCGCCG 251
Oy 236 CTCTCCCGAGACCGGGGGGGCTTTGGATTT-----GGGGGGGGGGGACGACTGCG 291
Db 252 CGCCGGACCATGTTCTGGGACCTGTACTTGGCGCTGCGCTGCTTGGACTGCTCT 311
```

```

Oy 292 CGGCGGACACATGTTCTTAAGCCACTCTGTACTTCGCGCTGCCACTCTTGATTTCTGAT 351
Db 312 GTGCGGCGGAAGTGAAGGAGGAGACCGGCTGGATTGGCGGAAGGACAGATCACTGCT 371
Oy 332 GTCCCGCCGAGGATGTGGAGACCGTGTGGACTGTGGAAAGCCAGGATCACTGCT 411
Db 372 GAAAGGACAGAGCTCAGCAGCCAGTACCGCAGCTAAAGGACAGTGTGGCGGCAAGA 431
Oy 412 GAAGGACAGAGCTCAGCAGCCAGTACCGCAGCTAAAGGACAGTGTGGCGGCAAGA 471
Db 432 GACCAACTTCAAGCTCGCATCCGCGCTGGAGCCAAAGATGATGATCCGCGCATGGA 491
Oy 472 AACCACACTTCAAGCTCGCATCCGCGCTGGAGCCAAAGATGATGATCCGCGCATGGA 531
Db 492 GGGCCCTGAAGCAGAGTGGCTTACAAACGCGGCTGCAAGCGGGGATGAAGAAGAGA 551
Oy 532 GCGCTTGAAGCAGAGTGGCTTCTGTACAACTGCGGCAAGCGGGGATGAAGAAGAGA 591
Db 552 GAACCTGCGCTGAGCTTACTGTGAGCATGTACAGAGCCGTGAGGAAATGATCTGCTGA 611
Oy 592 GAATGTCTGCGTATCTACTGTGAGCATGTACAGAGCCGTGAGGAAATGATCTGCTGA 651
Db 612 GGATTCGCCATATGAACAGATTTACAGAGATGTTCAGATATATTCGGGTGCCAT 671
Oy 652 AGATTCCCGGTATGAGCCGTTACAGAGGTTGTAGATATATTCGGGCGAGTCCGCT 711
Db 672 CATATCAG-TG-----G--AG-----CACATTCGCAAGGGAACAACTGCTGATGC 716
Oy 712 CATATCAGATGTGTTTCCAGCAAGTGGAAACATTTCCAAAGGGAACAACTGCTGATGC 771
Db 717 AGCGAAGGCTGCAACCTGAGCAGATTTGCAAGAGTACAGTGGCGTATCATCAACC 776
Oy 772 AGCGAAGGCTGCAACCTGAGCAGATTTGCAAGAGTACAGTGGCGTATCATCAACC 831
Db 777 GTGACACACAGCGTGTCCAGAGATGTCTGCAACGCGCCAAAGTCCACAGAGCCCTCG 836
Oy 832 CTGACACACAGCGTGTCCAGAGATGTCTGCAACGCGCCAAAGTCCACAGAGCCCTCG 891
Db 837 GGAGTCTTTGCAAGAGTCCCGGCAAGCAGTACGGAATGCTCTTGTCTCTGCGG 896
Oy 892 GGAGTCTTTGCAAGAGTCCCGGCAAGCAGTACGGAATGCTCTTGTCTCTGCGG 951
Db 897 GGAGTCTTTGCAAGAGTCCCGGCAAGCAGTACGGAATGCTCTTGTCTCTGCGG 956
Oy 952 GGAGTCTTTGCAAGAGTCCCGGCAAGCAGTACGGAATGCTCTTGTCTCTGCGG 1011
Db 957 GAGGAGAGCCCAACTGTTGAAATTTGCGAGACTCTCTGCAAGAGCAATTTACATCTGAG 1016
Oy 1012 ACGAGAGAGGCGCAACTGCTGAGTCTGCAAGACTCTCTGCAAGAGCAATTTACATCTGAG 1071
Db 1017 ATCTGCGCTTGGGATTTTTTTTACCACTGCGCAGAGCAAGTCAAGTCTGTGACAGCTG 1076
Oy 1072 ATCTGCGCTTGGGATTTTTTTTACCACTGCGCAGAGCAAGTCAAGTCTGTGACAGCTG 1131
Db 1077 TCTAAGGAGAACTAGCGTGTGAGTCTCTGCGCTACTGCGGCTTATTTGAGCAGTCA 1136
Oy 1132 TCTTAAAGGAACTAGCGTGTGAGTCTCTGCGCTACTGCGGCTTATTTGAGCAGTCA 1191
Db 1137 GACCCCACTACATAGCTCAGTACGCTCAGTGTGCGCCATGGTGTGACTGAGCAA 1196
Oy 1192 GACTCCCACTACATAGCTCAGTACGCTCAGTGTGCGCCATGGTGTGACTGAGCAA 1251
Db 1197 CAGTGGGAACGACTTGAAGAGTGTGAAATTTTGAATTTCTTCAAGGACAAATCATG 1256
Oy 1252 CAGCGGCACTGAGTGTGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1311
Db 1257 TCTTAAATTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1316
Oy 1312 TCTTAAATTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1371
Db 1317 CTTCCAGTACAGACCACTGCACTACACAGTCCCTCCGGGTTAAGAACAGCC 1376
Oy 1372 CCCTCCAGTACAGACCACTGCACTACACAGTCCCTCCGGGTTAAGAACAGCC 1431

```

```

Db 1377 CTTGGGCGACAGAGTCTGAGATGAATTTCCACTCATGTTTGGCACCCTGTGCAA 1436
Oy 1432 TCTGGGCGACAGAGTCTGAGATGAATTTCCACTCATGTTTGGCACCCTGTGCAA 1491
Db 1437 TTTACAGCAGAGAGTGAATTTCAATGTGTGGGCAATACACACTGTGTATTTCAA 1496
Oy 1492 TTTGAGGCTCAGAAAGTGAATTTCAATGTGTGGGCAATACACACTGTGTATTTGA 1551
Db 1497 TGTATTTATTAATAAAGAGTCTGCG---GTGCTTCCAGCCACATTAACCAATAATCAAT 1553
Oy 1552 TAGTATTTTGAAGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1611
Db 1554 GCGTCTCTCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1613
Oy 1612 GCGTCTCTCTCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1671
Db 1614 CACCTTATTTATTTTA 1629
Oy 1672 TGCCCTGTATCTGTA 1687

RESULT 11
LOCUS AF012811 777 bp mRNA ROD 06-AUG-1997
DEFINITION Mus musculus GDNF receptor alpha mRNA, partial cds.
ACCESSION AF012811
NID 92286224
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
1..777
/organism="Mus musculus"
/db_xref="taxon:10090"
1..430
/organism="Mus musculus"
/clone="402011"
/note="EST deposited under GenBank Accession Number W76716 from Soares mouse embryo NDME13.5 14.5 cDNA library; IMAGE Consortium"
<1..680
/codon_start=3
/product="GDNF receptor alpha"
/db_xref="PDB:1g2286225"
/translating="LNLQDSCKNYICRSLADFTNCPESRVSNCLEKNYADCL
AYSLGIVMTAPVYIDSSLSVAPWCDSNSGDLDECLFLNFKNDTCLKNAIOF
GNSGDTVMQDPAPVOTTTATTTAATFAIRKPKPGAPASENEIPIHVPICANQAOFL
KSNVSGSTHCLSDNDYGRKDLGASSHITTKMAAPSPSGLSLIPVPTALALIS
VSLAERS"
BASE COUNT 202 a 212 c 174 g 189 t
ORIGIN
Query Match 29.9% Score 640; DB 23; Length 777;
Best Local Similarity 93.6% Pred. No. 0.00e+00;
Matches 735; Conservative 0; Mismatches 41; Indels 9; Gaps 2;
Db 1 GCGTGAATCTGCAAGACTCTGCAAGAGCAATTTACATCTGAGATCTGCGCTTGAGATT 60
Oy 1029 GCGTAGTCTGCAAGACTCTGCAAGAGCAATTTACATCTGAGATCTGCGCTTGAGATT 1088
Db 61 TTTTACCAACTGCGCAGCAGAGTCAAGTCTGTACAGCAATGTCTTAAAGAGAACTAGC 120
Oy 1089 TTTTACCAACTGCGCAGCAGAGTCAAGTCTGTACAGCAATGTCTTAAAGAGAACTAGC 1148
Db 121 CAGACTGCTCTGCGCTACTGCGGACTGATTTGGACACAGTCACTATCTTAACATACATAG 180

```





QY 1028 TGCCTGAGCTGCGAAGACTCCTGCAAGACCAATTAATCTGAGATCTGCGCTTCAGAT 1087  
DB 1523 TTTTTCACAACTGCGAGCCTGAGTACAGCTGTAGTGTCTGTAAGAGAGAACTAC 1582  
QY 1088 TTTTTCACAACTGCGAGCCTGAGTACAGCTGTAGTGTCTGTAAGAGAACTAC 1147  
DB 1583 GCTGACTGCTCTGCTGCTTACTGAGGAGCTCATTTGGCAGACAGTACACCGAATACATA 1642  
QY 1148 GCAGACTGCTCTGCTGCTTACTGAGGAGCTCATTTGGCAGACAGTACACCGAATACATA 1207  
DB 1643 GACTCAAGTACTGCTGCTTACTGAGGAGCTCATTTGGCAGACAGTACACCGAATACATA 1702  
QY 1208 GACTCAGAGAGCTGAGCTGCTGCTTACTGAGGAGCTCATTTGGCAGACAGTACACCGAAT 1267  
DB 1703 GATGATGCGGGAATTTCTGTAATTTCTTCCAGSACCAACATGCTTAAATGCAAT 1762  
QY 1268 GAGAGCTGCTTAAATTTCTGTAATTTCTTCCAGSACCAACATGCTTAAATGCAAT 1327  
DB 1763 GAGGCTTTCGCAATGCTGCTGCTTACTGAGGAGCTCATTTGGCAGACAGTACACCGAAT 1822  
QY 1328 CAAGCTTTGGCAATGCTGCTGCTTACTGAGGAGCTCATTTGGCAGACAGTACACCGAAT 1387  
DB 1823 ACTACAGCCACACTACAGTACAGTTCAGACTTAAACACAGTTCAGACACCAAC 1882  
QY 1388 ACCACTGCCACCACTACAGTTCAGACTTAAACACAGTTCAGACACCAAC 1446  
DB 1883 AA-TGA-AAT-ACCAACCCACA-ATGATTCACAGAGTGGGAACTGCGAGACAGAA 1938  
QY 1447 GCTGAGATGAGATCCACACAGCTTTTACACCTGCTGGAATTTGACAGCTCAGAA 1506  
DB 1939 GAGCGGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1998  
QY 1507 G---GTGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563  
DB 1999 GAAAGA 2004  
QY 1564 AAGA 1569

RESULT 13  
LOCUS HSGFRA1G06 596 bp DNA PRI 10-APR-1998  
DEFINITION Homo sapiens GPI-linked anchor protein (GFRAL) gene, exon 6.  
ACCESSION AF038415  
NID 92921538  
KEYWORDS  
SEGMENT 6 of 11  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS Angrist,M., Jing,S., Bolk,S., Bentley,K., Nallasamy,S.,  
Halushka,M., Fox,G.M. and Chakravarti,A.  
TITLE Human GFRAL: Cloning, mapping, genomic structure, and evaluation as  
a candidate gene for Hirschsprung disease susceptibility  
JOURNAL Genomics 48, 354-362 (1998)  
REFERENCE  
AUTHORS Angrist,M., Jing,S., Bolk,S., Bentley,K., Nallasamy,S.,  
Halushka,M., Fox,G.M. and Chakravarti,A.  
TITLE Direct Submission  
JOURNAL Submitted (12-DEC-1997) Genetics, Case Western Reserve, 2109  
Adelbert Road, Cleveland, OH 44106-4955, USA  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..596  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q25"  
147..483  
/gene="GFRAL"  
exon  
BASE COUNT 152 a 156 c 163 g 125 t  
ORIGIN

Query Match 12.9%; Score 275; DB 22; Length 596;  
Best Local Similarity 90.6%; Pred. No. 1.19e-205;  
Matches 307; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

DB 145 AGTGAGACATTTCCCAAGGAACCACTGCTGATGACAGGAGGCTCAACTGA 204  
QY 733 AGTGAGACATTTCCCAAGGAACCACTGCTGATGACAGGAGGCTCAACTGA 792  
DB 205 CGACATTTGCAAGAGTACAGTGGGCTTACATCAACCCGTCACACACCGTGTCCA 264  
QY 793 CGACACCTGTAAGAGTACAGTGGGCTTACATCAACCCGTCACACACCGTGTCCA 852  
DB 265 CGATGCTGCAACCGCCGCAAGTGGCAGACAGGCTTCCGAGATTTCTTGAAGTCC 324  
QY 853 CGAGTCTGCAACCGCCGCAAGTGGCAGACAGGCTTCCGAGATTTCTTGAAGTCC 912  
DB 335 GGCCAAAGCAGCTACAGGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384  
QY 913 GGCCAAAGCAGCTACAGGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972  
DB 385 GAGCGACAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 444  
QY 973 GCGGCGACAGACTACGCTCCCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1032  
DB 445 GAATTTGCAAGCTCTGCAAGACGATTAATCATCTGACG 483  
QY 1033 GAGTGTGCAAGCTCTGCAAGACGATTAATCATCTGACG 1071

RESULT 14  
LOCUS HSGFRA1G03 497 bp DNA PRI 10-APR-1998  
DEFINITION Homo sapiens GPI-linked anchor protein (GFRAL) gene, exon 3.  
ACCESSION AF038412  
NID 92921535  
KEYWORDS  
SEGMENT 3 of 11  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS Angrist,M., Jing,S., Bolk,S., Bentley,K., Nallasamy,S.,  
Halushka,M., Fox,G.M. and Chakravarti,A.  
TITLE Human GFRAL: Cloning, mapping, genomic structure, and evaluation as  
a candidate gene for Hirschsprung disease susceptibility  
JOURNAL Genomics 48, 354-362 (1998)  
REFERENCE  
AUTHORS Angrist,M., Jing,S., Bolk,S., Bentley,K., Nallasamy,S.,  
Halushka,M., Fox,G.M. and Chakravarti,A.  
TITLE Direct Submission  
JOURNAL Submitted (12-DEC-1997) Genetics, Case Western Reserve, 2109  
Adelbert Road, Cleveland, OH 44106-4955, USA  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..497  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q25"  
76..369  
/gene="GFRAL"  
exon  
BASE COUNT 94 a 174 c 151 g 78 t  
ORIGIN

Query Match 11.4%; Score 244; DB 22; Length 497;  
Best Local Similarity 91.2%; Pred. No. 2.41e-178;  
Matches 270; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

DB 75 GACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 134  
QY 341 GATTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400





\*\*\*\*\*  
 WIRE RELEASE (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsearch n.a. - n.a. database search, using Smith-Waterman algorithm

on: Wed Jun 24 23:10:17 1998; Maspar time 314.60 Seconds  
 865.544 Million cell updates/sec

Output not generated.

Title: >US-08-866-354-3  
 Description: (1-2138) from US08866354.seq  
 Perfect Score: 2138  
 N.A. Sequence: 1 AGCTGGCTCTCCCGGGGCGAG.....AGACAGCGCCGAGCGCTGC 2138  
 Comp: TCGACCGAGAGGGGCGCCGCTC.....TCGTCCGGCGCTCGGCGAC

Scoring table: TABLE default  
 Gap 6

Mmatch STD : DBase 0; Query 0

Searched: 176886 seqs, 63680241 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database:

n-geneseq31-2  
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
 14:part14 15:part15 16:part16 17:part17 18:part18  
 19:part19 20:part20 21:part21 22:part22 23:part23  
 24:part24 25:part25 26:part26 27:part27 28:part28  
 29:part29 30:part30 31:part31 32:part32 33:part33  
 34:part34 35:part35 36:part36 37:part37

Statistics: Mean 9.707; Variance 5.482; scale 1.771

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1885	88.2	2378	37	Rat glial cell derive	0.00e+00
2	748	35.0	840	37	Mouse GDNF alpha clo	0.00e+00
3	347	16.2	418	37	Mouse GDNF alpha clo	2.00e-230
4	133	6.2	453	37	Human EST-derived seq	1.41e-72
5	122	5.7	351	37	Human EST-derived seq	1.01e-64
6	112	5.2	385	15	DNA probe 32 detects	1.25e-57
7	96	4.5	201	37	Human EST-derived seq	2.10e-46
8	96	4.5	201	37	Human EST-derived seq	2.10e-46
9	44	2.1	91	9	Oligonucleotide probe	7.63e-12
10	44	2.1	91	9	Oligonucleotide probe	7.63e-12
11	43	2.0	204	1	Base substituted E.co	3.11e-11
12	40	1.9	204	1	Base substituted E.co	1.99e-09
13	38	1.8	114	12	Generic DNA sequence	3.05e-08
14	38	1.8	114	12	Generic DNA sequence	3.05e-08

15	38	1.8	114	12	Generic DNA sequence	3.05e-08
16	36	1.8	172	32	Human interleukin 8 a	7.83e-09
17	39	1.7	114	12	Generic DNA sequence	4.46e-07
18	36	1.7	114	12	Generic DNA sequence	4.46e-07
19	36	1.7	190	31	Chymase antisense o11	6.22e-06
20	34	1.6	114	12	Generic DNA sequence	6.22e-06
21	34	1.6	114	12	Generic DNA sequence	6.22e-06
22	35	1.6	114	12	Generic DNA sequence	1.67e-06
23	34	1.6	114	12	Generic DNA sequence	6.22e-06
24	34	1.6	114	12	Generic DNA sequence	6.22e-06
25	34	1.6	114	12	Generic DNA sequence	6.22e-06
26	33	1.5	114	12	Generic DNA sequence	2.28e-05
27	33	1.5	114	12	Generic DNA sequence	2.28e-05
28	33	1.5	114	12	Generic DNA sequence	2.28e-05
29	32	1.5	114	12	Generic DNA sequence	8.22e-05
30	32	1.5	168	31	Human MDRF antisense	8.22e-05
31	33	1.5	172	31	Human interleukin 8 a	2.28e-05
32	32	1.5	178	31	Human interleukin-1 an	8.22e-05
33	29	1.4	42	3	MCPC 603 VH CDR2 walk	3.52e-03
34	30	1.4	100	31	Human IL4 receptor an	1.02e-03
35	29	1.4	114	12	Generic DNA sequence	3.52e-03
36	30	1.4	178	31	Human endothelin-1 an	1.02e-03
37	31	1.4	250	31	Substance P antisense	2.92e-04
38	31	1.4	3871	2	HSV-1 g8 and surround	2.92e-04
39	27	1.3	39	7	Mixed oligonucleotide	3.95e-02
40	27	1.3	114	12	Generic DNA sequence	3.95e-02
41	27	1.3	128	31	Human IL6 antisense o	3.95e-02
42	28	1.3	162	31	Human RANTES antisens	1.19e-02
43	28	1.3	190	31	Chymase antisense o11	1.19e-02
44	28	1.3	264	31	Substance P receptor	1.19e-02
45	27	1.3	264	31	Substance P receptor	3.95e-02

## ALIGNMENTS

RESULT 1  
 ID T84975 standard; cDNA; 2378 BP.  
 AC T84975:  
 DT 27-APR-1998 (first entry)  
 DE Rat glial cell derived neurotrophic factor receptor alpha cDNA.  
 KW Glial cell derived neurotrophic factor receptor alpha; GDNF alpha;  
 KW GDNF; rat; kidney disease; glomerulonephritis; therapy; ds.  
 OS Rattus sp.  
 FH Key  
 FT CDS location/Qualifiers  
 FT sig\_peptide /\*tag= a  
 FT /\*tag= b  
 FT mat\_peptide 189..1520 /\*tag= c  
 FT W09733912-A2.  
 PD 18-SEP-1997.  
 PR 13-MAR-1997; U04363.  
 PR 14-MAR-1996; US-618236.  
 PR 14-MAR-1996; US-615902.  
 PA (GETH ) GENENTECH INC.  
 PI Klein RD, Moore MW, Rosenthal A, Ryan AM;  
 DR WPI: 97-470819/43.  
 DR P-PSDB: W27327.  
 PT Isolated glial cell derived neurotrophic factor receptor alpha -  
 PT useful to develop products to diagnose and treat associated  
 PT disorders, particularly enteric nervous system or kidney disorders  
 PS Clam 23; Page 74-77; 100pp; English.  
 CC This cDNA codes for full-length rat glial cell derived neurotrophic  
 CC factor receptor alpha (GDNF alpha) (see W27327), a novel  
 CC GPI-linked protein that is a ligand-binding component of the  
 CC receptor system for GDNF. It was isolated by expression cloning.  
 CC A cDNA library obtained from ventral midbrain tissue of E14 rat  
 CC embryos was generated in a cyomegalovirus-based vector. cDNA  
 CC clones were transfected into COS 7 cells and expression of  
 CC putative GDNF receptors was detected by binding of iodinated GDNF.  
 CC A single positive pool was obtained, from which the cDNA clone was  
 CC isolated. An expression vector containing the cDNA can be used to

CC produce GDNF alpha in transformed host cells. The invention  
CC relates to novel uses of GDNF and its receptor. In particular, it  
CC relates to native rat GDNF alpha (see W27327), its variants and  
CC soluble derivatives (extracellular domain), chimeric GDNF alpha  
CC and antibodies which bind to the GDNF alpha, including agonist  
CC and neutralizing antibodies, as well as various uses for these  
CC molecules. It also relates to assay systems for detecting ligands  
CC to GDNF alpha, systems for studying the physiological role of  
CC GDNF, diagnostic techniques for identifying GDNF-related conditions,  
CC methods for identifying molecules homologous to GDNF alpha, and  
CC therapeutic techniques (claimed) for the treatment of GDNF-related  
CC and GDNF alpha-related conditions, particularly kidney disease  
CC associated with glomerulonephritis and enteric nervous system  
CC related disorders. Transgenic and knockout animals are also  
CC claimed.  
SQ Sequence 2378 BP: 617 A: 624 C: 598 G: 539 T:  
Query Match 88.2%; Score 1885; DB 37; Length 2378;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1886; Conservative 0; Mismatches 1; Indels 0; Gaps 0:  
DB 63 ccggcgccggtcttgatcttggggggcggggaacagctcgccgcgccacatgttc 122  
|||  
QY 248 CCGGGCGGGCGCTTGGATTGTTGGGGGGCGGGACCAAGCTCGCGGGCACCATGTTC 307  
|||  
DB 123 ctggcaactctgactctgacgctgacactctgatttgcgactgctcgccgagtgagt 182  
|||  
QY 308 CTAGCCACTCTGTACTTGCCTGCGCTGCACTCTGTGATTTGCTGATGTCGCCGAGATGAGT 367  
|||  
DB 183 ggtggaagccgctcgtgactgtgtgaagccagcgatcagtcctggaagaaacagagctgc 242  
|||  
QY 368 GGTGAGACCGCTGTGACTGTGTAAAGCCAGCATGTCCTGGAAGAAAGAGAGAGTGC 427  
|||  
DB 243 agcccaagtcgcccaactaagcagtgctggtcgggcaagaaacacacttcaagctg 302  
|||  
QY 428 AGCACAAGTACCCACACTAAGCAGTGTGCGGGCAAGGAACCACTTACGCTG 487  
|||  
DB 303 acatccggcctctgagggccaaagagatgagtcgctagcgccatggagggccttgaacagaag 362  
|||  
QY 488 ACATCCGGCTTGAAGCCCAAGATGATGTCGTAAGCCGATGAGAGGCTTAAACAGAGAG 547  
|||  
DB 363 tctctgtaacaactgcgctgcaagcggggcaatgaagaagaagaatctgctgctatc 422  
|||  
QY 548 TCTCTGTATCAACTGCGCTGCAGCGGGGCAATGAAGAAAGAAATGTGTCTCGTATC 607  
|||  
DB 423 tactggagcatgtaccagagctctgcaaggaaatgaactcctggaagattcccgatagag 482  
|||  
QY 608 TACTGGAGCATGTACAGAGCTGCAGGAATGACCTCTGGAAGATTTCCCTATGAG 667  
|||  
DB 483 ccggtttaaagcaggttgcagatatctcgggcagctccgcttcatactcagatgttttc 542  
|||  
QY 668 CCGGTAAACACAGAGTTGTGATATATTCGGGCAAGTCCGTTCAATATCAAGATTTTC 727  
|||  
DB 543 cagaagaatggaacacatttccaagaaggaaacactgctggaacgcaagaagctctgaac 602  
|||  
QY 728 CAGCAAGTGGAAACATTTCAAAGGGAACACTGCTGGAGGAGCAAGCAAGCTTGCAAC 787  
|||  
DB 603 ctggaagcagacctgttaagaagtacaggtcggtctacatcacccctgcaacacagatg 662  
|||  
QY 788 CTGACACACACTCTAAGAAATACAGGTGCGCTTACATCACCCCTGACACACAGCATG 847  
|||  
DB 663 tccaagaggtctgcaaccgctgaagtgcgaagaagccctcagaagcaggttcttgaagaag 722  
|||  
QY 848 TCNAAGAGGTCTCAACCGCTGTAAGTGCACAAAGGCTTCAGAGAGTTTGTGACAAAG 907  
|||  
DB 723 gttccggccaaagcagctaacggagatgctctctgctcctcgaacacatcgctgcaac 782  
|||  
QY 908 GTTCCGGCAAGCAAGCATGTAAGGATGCTCTTGTCTCTCTCCGGAGCATGCTTGCAAC 967  
|||  
DB 783 gaggcgcgcgagacagactatcgtcccggtgtgtcctatgaagaacgagagggccaac 842  
|||  
QY 968 GAGCGGGGCGACAGACTATGTCCTCGGTGTCTATGAAGAAAGAGAGGCCCAAC 1027  
|||

DB 843 tgcctgagctgcaagaagactcctctcaagaccatatactcagagatcgtccgacagat 902  
|||  
QY 1028 TGCTGTAGTGTGCAAGACTCTCTGCAAGACCAATTAATCTGCGAGATCTGCCCTTGCAAGT 1087  
|||  
DB 903 ttttttcaactctgcaagcagagatcgaaggtctgtcagcaactgtcttgaagaaactac 962  
|||  
QY 1088 TTTTATCAACTGCGCAGCCAGAGTCAAGGTCTGTGAGCAACTGTCTTAAGAGAACTAC 1147  
|||  
DB 963 gcaagcgtcctctgctcctcctcctcgtggaactatgtgcaagtcatactgaactcccaactga 1022  
|||  
QY 1148 GCAAGACTGCTCTGTGGCTCTCTGCGGCTATTTGGCACTATGACTCCCACTACAGTGA 1207  
|||  
DB 1023 gactccagagcctcagcgtggaacacatgtgtgactgcaagcagcagcgaatgacgtg 1082  
|||  
QY 1208 GACTCCAGAGCTCTACAGCTGCGACCATGCTGTGACTGCGCAACAGCGCAATGACTG 1267  
|||  
DB 1083 gaagactgtcttgaatctctgaattttttaaaggaacatactgtctcaaaaatgcaat 1142  
|||  
QY 1268 GAAGACTGCTTGAATAATTTGAAATTTTAAAGACAAATCTGTCTCAAAATGCAAT 1327  
|||  
DB 1143 caagccttggcaatgctcagatgagacatgtggaagcagcagccctccagtcagacc 1202  
|||  
QY 1328 CAAGCCTTTGGCAATGGCTCAGATGTGACCATGTGTGACAGCCACCCCTCCAGTCCAGACC 1387  
|||  
DB 1203 acaactgccacaactacactgctcctcgggtcagaagaacagcctctggggccagcaggg 1262  
|||  
QY 1388 ACCACTGCCACACTACACTGCTCTCCGGGTCAAGAAACAGCCTCTGGGGCCAGCAGGG 1447  
|||  
DB 1263 tctgaagaatgagatccccaacacagcttllaacacccctgtggaatttgcagagctcagaag 1322  
|||  
QY 1448 TGTGAGAAATAGATCCCAACACAGCTTTACACACCTGTGCAATTTGACAGGCTCAGAG 1507  
|||  
DB 1323 ctgaataccaaatgtgtcgggtgaacacacacactctgtctctgaatgtgttccggaag 1382  
|||  
QY 1508 CTGAATTCAAATGTGTGCGGTAGCACACACCTCTGTTCTATATGTGATTTGCGAAG 1567  
|||  
DB 1383 gatgtctcgtctgtgctcagcacaacacacacacacacacacacacacacacacacacac 1442  
|||  
QY 1568 GATGCTTCCCTGCTGCTGCTCAGCCACATTAACCAAAATCAATGCTCTCTCCAGC 1627  
|||  
DB 1443 tgcagctcagactcactgcggtgtcgtatgtcagccgctctgctcctctgtatctgta 1502  
|||  
QY 1628 TGAGTGTGAGTCACTAGCGGCTGTGATGCTCAGCGCCCTGTGCTGCTTATGTGTA 1687  
|||  
DB 1503 tctgttgcaagaacgtcgtatgcgtcatccgggaagaacagatgaagaagaagaagaac 1562  
|||  
QY 1688 TCGTTGGCAAAACGTCGTAGCTGCAACCCGGGAACAGATGATAAAGACAAAAGAGAAC 1747  
|||  
DB 1563 caagtatctgtccctgtcctcctctgtatctcgaaaatccagctttaaagaagctcgctga 1622  
|||  
QY 1748 CAAGTATTTCTGCTCCTGCTCTTGTATATCTGAAAATCCAGTTTAAAGCTCCGTTGA 1807  
|||  
DB 1623 gaagcagtttcaacccaactggaactcttctgttctttaaagaagcctgtgtgcccctcag 1682  
|||  
QY 1808 GAAGCACTTTCACCAACATGGAACCTTCTCTGTTTAAAGAAAGTTGTGCCCCCTCAG 1867  
|||  
DB 1683 gggctctgttgaagaacgtctacagggcctaattccaacccaatgaagctctggggcgtg 1742  
|||  
QY 1868 GGGCTTGTGTGAAGAACTCTACAGGGCTTAATTCMAAACCACTAAGGCTCTGGGGGCTG 1927  
|||  
DB 1743 gtcggccttgaaggggacacttgcacacatgtaaagaagcttggtgctatcatgtgttga 1802  
|||  
QY 1928 GTGGCGCTTAAAGGGGACATTTTCACCATATTAAGCAAGCTGGGCTTATATGTGTTGA 1987  
|||  
DB 1803 tggtagaagatgtagttagttagttagttagttagttagttagttagttagttagttagttagt 1862  
|||  
QY 1988 TGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2047  
|||  
DB 1863 tactgttagaagcaggaagatatactatgaagaatcttcaatgtcttactcagacga 1922  
|||  
QY 2048 TACTGTTAGGAACAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2107  
|||  
DB 1923 ttcgcttctgaagaagcgcccgagcc 1949  
|||

OY 2108 TTGCTTCTGAAGACAGCGCCGACGCC 2134  
 RESULT 2  
 ID T84977 standard; cDNA: 840 BP.  
 AC T84977;  
 DT 27-APR-1998 (first entry)  
 DE Mouse GDNFR alpha clone 26 3' end.  
 KW GDNFR alpha derived neurotrophic factor receptor alpha; GDNFR alpha;  
 KM GDNFR; mouse; kidney disease; glomerulonephritis; therapy; ss.  
 OS Mus musculus.  
 PN W09733912-A2.  
 PD 18-SEP-1997.  
 PR 13-MAR-1997; U04363.  
 PR 14-MAR-1996; US-618236.  
 PR 14-MAR-1996; US-615902.  
 PA (GETH ) GENENTECH INC.  
 PI Klein RD, Moore MW, Rosenthal A, Ryan AM;  
 WPI: 97-470819/43.  
 Isolated glial cell derived neurotrophic factor receptor alpha -  
 useful to develop products to diagnose and treat associated  
 disorders, particularly enteric nervous system or kidney disorders  
 Example 1; Page 60-61; 100pp; English.  
 CC This cDNA sequence comprises the 3' end of mouse full-length  
 CC glial cell derived neurotrophic factor receptor alpha (GDNFR alpha)  
 CC (see W27327) clone 26; the 5' end of the clone is given in  
 CC T84978. Clone 26 was isolated from a mouse cDNA library using  
 CC rat GDNFR alpha cDNA (see T84975) as probe. The invention relates  
 CC to novel uses of GDNFR and its receptor. In particular, it relates  
 CC to native rat GDNFR alpha (see W27327), its variants and soluble  
 CC derivatives (extracellular domain), chimeric GDNFR alpha and  
 CC antibodies which bind to the GDNFR alpha, including agonist and  
 CC neutralising antibodies, as well as various uses for these  
 CC molecules. It also relates to assay systems for detecting ligands  
 CC to GDNFR alpha, systems for studying the physiological role of  
 CC GDNFR, diagnostic techniques for identifying GDNFR-related conditions,  
 CC methods for identifying molecules homologous to GDNFR alpha, and  
 CC therapeutic techniques (claimed) for the treatment of GDNFR-related  
 CC and GDNFR alpha-related conditions, particularly kidney disease  
 CC associated with glomerulonephritis and enteric nervous system  
 CC related disorders. Transgenic and knockout animals are also  
 CC claimed.  
 CC Sequence 840 BP: 205 A: 245 C: 200 G: 190 T:  
 SO  
 Query Match 35.0%; Score 748; DB 37; Length 840;  
 Best Local Similarity 94.5%; Pred. No. 0.00e+00;  
 Matches 794; Conservative 0; Mismatches 46; Indels 0; Gaps 0;  
 1 CGCGCAAGTGCACAAAGCCCTCAGGCAAGTTCCTGCAAGAGTCCGACGACGAGC 60  
 OY CCCCCCTAAGTGCACAAAGCCCTCAGGCAAGTTCCTGCAAGAGTCCGACGACGAGC 925  
 DB 61 TACGGAGATCCTCTCTGCTCCGCGGAGCGTCCGCTGACGAGAGGAGGAGCAAGACT 120  
 OY TACGGAGATCCTCTCTGCTCCGCGGAGCGTCCGCTGACGAGAGGAGGAGCAAGACT 985  
 DB 121 ATCGTCCCTGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 180  
 OY ATCGTCCCTGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1045  
 DB 181 TCTGTCAAGACAAATTAATCTGCGAGATCTGCGCTTGTGCAAGTCTGCAAGTCCGAG 240  
 OY TCTGTCAAGACAAATTAATCTGCGAGATCTGCGCTTGTGCAAGTCTGCAAGTCCGAG 1105  
 DB 241 CCAGAGTCAAGTCTGCTGCGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 OY CCAGAGTCAAGTCTGCTGCGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1165  
 DB 301 TACTCGGAGATCTGCTGCGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 OY TACTCGGAGATCTGCTGCGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1225

DB 361 GTGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 OY GTGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1285  
 DB 421 CTGAATTTTAAAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 OY CTGAATTTTAAAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345  
 DB 481 TCGATGTGACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 OY TCGATGTGACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1405  
 DB 541 ACTGCTTCCGCGTCAAG 600  
 OY ACTGCTTCCGCGTCAAG 1465  
 DB 601 ACACAGCTTTTACCAAG 660  
 OY ACACAGCTTTTACCAAG 1525  
 DB 661 GCGAGTACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 OY GCGAGTACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1585  
 DB 721 TCCAGCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 780  
 OY TCCAGCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1645  
 DB 781 CCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 OY CCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1705  
 RESULT 3  
 ID T84976 standard; cDNA: 418 BP.  
 AC T84976;  
 DT 27-APR-1998 (first entry)  
 DE Mouse GDNFR alpha clone 26 5' end.  
 KW GDNFR alpha derived neurotrophic factor receptor alpha; GDNFR alpha;  
 KM GDNFR; mouse; kidney disease; glomerulonephritis; therapy; ss.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT CDS 287..418  
 FT /note="5' end of GDNFR alpha coding region"  
 FT  
 PN W09733912-A2.  
 PD 18-SEP-1997.  
 PR 13-MAR-1997; U04363.  
 PR 14-MAR-1996; US-618236.  
 PR 14-MAR-1996; US-615902.  
 PA (GETH ) GENENTECH INC.  
 PI Klein RD, Moore MW, Rosenthal A, Ryan AM;  
 WPI: 97-470819/43.  
 Isolated glial cell derived neurotrophic factor receptor alpha -  
 useful to develop products to diagnose and treat associated  
 disorders, particularly enteric nervous system or kidney disorders  
 Example 1; Page 60; 100pp; English.  
 CC This cDNA sequence comprises the 5' end of mouse full-length  
 CC glial cell derived neurotrophic factor receptor alpha (GDNFR alpha)  
 CC (see W27327) clone 26; the 3' end of the clone is given in  
 CC T84977. Clone 26 was isolated from a mouse cDNA library using  
 CC rat GDNFR alpha cDNA (see T84975) as probe. The invention relates  
 CC to novel uses of GDNFR and its receptor. In particular, it relates  
 CC to native rat GDNFR alpha (see W27327), its variants and soluble  
 CC derivatives (extracellular domain), chimeric GDNFR alpha and  
 CC antibodies which bind to the GDNFR alpha, including agonist and  
 CC neutralising antibodies, as well as various uses for these  
 CC molecules. It also relates to assay systems for detecting ligands  
 CC to GDNFR alpha, systems for studying the physiological role of  
 CC GDNFR, diagnostic techniques for identifying GDNFR-related conditions,  
 CC methods for identifying molecules homologous to GDNFR alpha, and  
 CC therapeutic techniques (claimed) for the treatment of GDNFR-related  
 CC and GDNFR alpha-related conditions, particularly kidney disease

associated with glomerulonephritis and enteric nervous system related disorders. Transgenic and knockout animals are also claimed.

Sequence 418 BP; 70 A; 119 C; 150 G; 79 T.

Query Match 16.2%; Score 347; DB 37; Length 418;  
Best Local Similarity 95.3%; Pred. No. 2,00e-230;  
Matches 365; Conservative 0; Mismatches 18; Indels 0; Gaps 0.

D 36 ggcaggttgagtcggaactgaacccttaaaagcggttcgcctccgcctcgcgcgcg 95  
OY 51 GGCAAGTTTGGGCGGAACTGAACCCCTGAAGCGGGTCCGCTCCGGCCTCGGCCCG 110

D 96 tcggagctgaatctccctgctggcgcggttggcgcgagcaagcaagggagttctctcaacct 155  
OY 111 CCGGATTCGAGCTCCGTGGCGGGCGGGCGGCAAGCAGCGGAGAGCTGCTTCACACCT 170

D 156 ggaatgagcttaactttagtcgcagagagggcgcaatcgcgcggggatctctcaagct 215  
OY 171 GGATGGAGACTGAACCTTTAGTGTGCGCAAGAGGCCAGTCGCCCGGAGTCCCTACAGCT 230

D 216 gagctctccgcgagatccggttgcgagctttgattttggggggcgagggagcaagctgc 275  
OY 221 GAGCTCTCTCCCGAGACCGGGCGGGCTTTGGATTTTGGGGGGCGGGAGCAGAGCTGC 290

D 276 gcggtgagcaatcttctctagcaactctgtactctctgctgcacactcctgattctga 335  
OY 291 GCGGCGGCAACATGTTCTTACCCACTCTGTACTTCGCGCTCCACTCTGGATTGCTGA 350

D 336 tctcgccgaggtgagtggtggggagccgcgagactgtgtaaagcaagtgatagcgcc 395  
OY 351 TGTCCGCGAGGTGAGTGTGTGAAGCCCTCTGAGACTGTGTAAAGCAGCGATAGTGC 410

D 396 tgaaggaaacagagctgcagcacc 418  
OY 411 TGAAGGAACAGAGCTGCAGCAC 433

RESULT 4  
ID T84979 standard; DNA; 453 BP.  
AC T84979;  
DT 27-APR-1998 (first entry)  
DE Human EST-derived sequence y170a10.r1.  
KW Glial cell derived neurotrophic factor receptor alpha; GDNFR alpha;  
KW GDNF; human; kidney disease; glomerulonephritis; therapy;  
KW expressed sequence tag; EST; ss.  
OS Homo sapiens.  
MO M09733912-A2.  
PI 18-SEP-1997.  
PR 13-MAR-1997; U04363.  
PR 14-MAR-1996; U5-618236.  
PR 14-MAR-1996; U5-615902.  
PA (GETH ) GENENTECH INC.  
PI Klein RD, Moore MM, Rosenthal A, Ryan AM;  
WP1: 97-470819/43.  
PT Isolated glial cell derived neurotrophic factor receptor alpha -  
PT useful to develop products to diagnose and treat associated  
PT disorders, particularly enteric nervous system or kidney disorders  
PS Example 1; Page 61; 100bp; English.  
CC This DNA sequence comprises a human EST-derived sequence designated  
CC y170a10.r1. It can be used to identify glial cell derived  
CC neurotrophic factor receptor (GDNFR) sequences, including human  
CC variants. Also suitable as probes for GDNFR sequences are another  
CC human EST-derived sequence designated y683h05.r1 (see T84978), and  
CC fragments of these EST-derived sequences (see T84980-81) or  
CC proteins encoded by them. The invention relates to novel uses of  
CC GDNF and its receptor. In particular, it relates to native rat  
CC GDNFR alpha (see W27327), its variants, and soluble derivatives  
CC (extracellular domain), chimeric GDNFR alpha and antibodies which  
CC bind to the GDNFR alpha, including agonist and neutralising  
CC antibodies, as well as various uses for these molecules. It also  
CC relates to assay systems for detecting ligands to GDNFR alpha.  
CC systems for studying the physiological role of GDNF, diagnostic

[illegible]









```
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcripiase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80375.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 2.0%; Score 43; DB 1; Length 204;
Best Local Similarity 22.1%; Pred. No. 3.1le-11;
Matches 36; Conservative 62; Mismatches 64; Indels 1; Gaps 1;

Dc 23 gcyyrcayccbcgagagycacbcyrragnyccccgggywcgcagcycaaycdchng 82
Cc 542 GCTTCACAGCGCTCATGCGCGTACGGCACCATCTCGTGCGCTCAAGGCCGAGT-CAGC 484
Cc 83 ccgyumrttthyrmdbvnyrdynrsdaawccyrirrvskvydcynachddhybbv 142
Cc 483 CTGAAGTTGCTTCTTCCTCCCGCCACGCACGCACGTCTGATTGCGCGTACTGTCTCAG 424
Db 143 ynvhnmcnccebnhcnvbnbnhnwyvhdariddvnh 185
Cc 423 CTCCTTCTTCACGACTGATCGCTGCTTCACACAGTCCA 381

RESULT 12
ID ID N81164 standard; DNM; 204 BP.
AC AC 08-NOV-1990 (first entry)
DT DT 08-NOV-1990 (first entry)
DE DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS OS Escherichia coli.
FH FH key location/Qualifiers
FT FT misc-feature 19..69
FT FT /*tag- a /function-multiple cloning site
FT FT primer_bind 187..204
PN PN EP-285123-A.
PD PD 05-MAY-1988.
PF PF 30-MAR-1988; 105163.
PR PR 03-APR-1987; US-034819.
PA PA (SUSO) SUOMEN SOKERI OY.
PI PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinkainen T;
PR PR WPI; 88-279927/40.
PP PP Introducing random point mutations into nucleic acids -
PP by prepn of single stranded template, annealing a primer, elongation,
PP misincorporation, completion of molecules and screening.
PS PS Disclosure: P; English.
CC CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80375.
SO Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 1.9%; Score 40; DB 1; Length 204;
Best Local Similarity 11.7%; Pred. No. 1.99e-09;
Matches 12; Conservative 55; Mismatches 35; Indels 1; Gaps 1;

Dc 85 gymtttthyrmdbvnyrdynrsdaawccyrirrvskvydcynachdddy-ybbdv 143
Cc 667 GCCGGTTAACAGCAGGTTGTGACGATATATTTCCGGGCACTCCGTTCAATCACAGTATTTT 726
```

[illegible]

ID 070465 standard; DNA; 114 BP.  
AC 070465:  
DT 05-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KM direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc-feature 55..60  
FT /\*tag- a  
FT /note- "this sequence represents 'Z'; Z can be a  
FT sequence of 6, 9 or 12 nucleotides (see  
FT comments)"  
PN MO9418318-A.  
PD 18-AUG-1994.  
PF 01-FEB-1994; U00977.  
PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PI (UYNC-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI; 94-279739/34.  
DR P-PSDB; R65150 and R65151.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins  
PT comprising a binding domain and an effector domain  
PS Disclosure; Page 35; 255pp; English.  
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
CC represented as follows: X(NNB)6(TGC)(NNB)14(TGC)(NNB)3Y. X  
CC and Y are flanking restriction sites (X is not the same as Y) that are  
CC not specified further. Other generic sequences are shown in Q70465-68.  
CC Other specific peptides generated by these generic sequences are shown in  
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
CC comprising at least two functional regions - a binding domain with  
CC affinity for a ligand and a second effector peptide portion that is  
CC chemically or biologically active. They may further comprise a linker  
CC peptide between the 2 domains. The oligonucleotides are also designed so  
CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
CC in, or flanking, the unpredicted or variant residues. These residues  
CC confer some degree of conformational rigidity to the peptides. The TSARs  
CC or compsns. comprising a TSAR binding domain can be used in vivo to  
CC deliver a chemically or biologically active moiety, eg. metal ion, to  
CC a radioisotope, peptide, toxin or enzyme, to the specific target or on the  
CC cell. They can also replace the function of macromolecules, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need  
CC for complex methods of hybridoma formation or in vivo antibody  
CC production. The TSARs are easily characterised and have designed  
CC activity allowing direct and rapid detection in a screening process.  
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;  
Query Match 1.88; Score 38; DB 12; Length 114;  
Best Local Similarity 4.58; Pred. No. 3.05e-08;  
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;  
Db 3 bnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnnnnnn 62  
QY 844 CAGTCCACGAGGTGTGCAACCGCCGTAAAGTCCACACAGGCCCTCAGGACGTTCTTGA 903  
DB 63 bnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnnb 114  
QY 904 CAGGTTCCGCCCAAGCAGCAGCTACGGAGTCTCTCTCTCTCCGCCGAGC 955  
RESULT 15  
ID 070469 standard; DNA; 114 BP.  
AC 070469:  
DT 07-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KM direct; rapid; detection; screening; treatment; generic; ss.

OS Synthetic.  
FH Key Location/Qualifiers  
FT misc-feature 55..60  
FT /\*tag- a  
FT /note- "this sequence represents 'Z'; Z can be a  
FT sequence of 6, 9 or 12 nucleotides (see  
FT comments)"  
PN MO9418318-A.  
PD 18-AUG-1994.  
PF 01-FEB-1994; U00977.  
PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PI (UYNC-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI; 94-279739/34.  
DR P-PSDB; R65150 and R65151.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins  
PT comprising a binding domain and an effector domain  
PS Disclosure; Page 35; 255pp; English.  
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide  
CC This generic formula can be represented as follows: X(TGC)(NNB)10-  
CC (TGC)(NNB)62(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction  
CC sites (X is not the same as Y) that are not specified further. This  
CC sequence generates peptides that are cloverleaf in structure. Other  
CC generic sequences are shown in Q70465-68. Other specific peptides  
CC generated by these generic sequences are shown in R65150-54. TSARs are  
CC concatenated heterofunctional proteins or peptides, comprising at least  
CC two functional regions - a binding domain with affinity for a ligand and  
CC a second effector peptide portion that is chemically or biologically  
CC active. They may further comprise a linker peptide between the 2 domains.  
CC The oligonucleotides are also designed so that the expressed peptide  
CC contains 2 or 4 cysteine residues positioned in, or flanking, the  
CC unpredicted or variant residues. These residues confer some degree of  
CC conformational rigidity to the peptides. The TSARs or compsns. comprising  
CC a TSAR binding domain can be used in vivo to deliver a chemically or  
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin  
CC or enzyme, to the specific target or on the cell. They can also replace  
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies  
CC and therefore circumvent the need for complex methods of hybridoma  
CC formation or in vivo antibody production. The TSARs are easily  
CC characterised and have designed activity allowing direct and rapid  
CC detection in a screening process.  
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;  
Query Match 1.88; Score 38; DB 12; Length 114;  
Best Local Similarity 7.18; Pred. No. 3.05e-08;  
Matches 8; Conservative 32; Mismatches 72; Indels 0; Gaps 0;  
Db 1 tgcdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnnnnnn 60  
QY 833 TGCAACCAAGCATGTCTCAACGAGGTGTGCAACCGCCGTAAAGTCCACAGGCCCTTAGG 892  
DB 61 nbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnnb 112  
QY 893 CAGTCTTCGACCAAGTTCGCCGACACAGCTACGGAGTCTCTCTCTCTCTCTCTCTCTCT 944  
Search completed: Wed Jun 24 23:20:17 1998  
Job time : 600 secs.

\*\*\*\*\*  
 WIREIMAGE (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPearch\_n n.a. - n.a. database search, using Smith-Waterman algorithm  
 on: Wed Jun 24 22:01:59 1998: Msearch time 2326.45 Seconds  
 1227.537 Million cell updates/sec  
 Tabular output not generated.

Title: >US-08-866-354-3  
 Description: (1-2138) from US08866354.seq  
 Perfect Score: 2138  
 N.A. Sequence: 1 AGCTGCGTCTCCCGGGGCGAG.....AGACAGCCCGCCGCGTCG 2138  
 Comp: TCGACGAGAGGGGCCCCGTC.....TCTGTCGCGGCGTCCGCGAC

Scoring table: TABLE default  
 Gap 6

Match STD : DBase 0; Query 0

Searched: 1759237 seqs, 667866413 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

## Database:

Database: emb1-est154  
 1:em-est1 2:em-est3  
 genbank-est106  
 3:gb-est1 4:gb-est10 5:gb-est11 6:gb-est12 7:gb-est13  
 8:gb-est14 9:gb-est15 10:gb-est16 11:gb-est17  
 12:gb-est18 13:gb-est19 14:gb-est2 15:gb-est20  
 16:gb-est21 17:gb-est22 18:gb-est3 19:gb-est4 20:gb-est5  
 21:gb-est6 22:gb-est7 23:gb-est8 24:gb-est9 25:gb-est  
 26:gb-est

Statistics: Mean 11.800; Variance 1.984; scale 5.947

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	466	23.2	631	7	AA518362	v110602.r1 Barstead mo
2	408	19.1	466	5	AA387098	vc18b07.r1 KO mouse em
3	401	18.8	465	21	W53498	md45601.r1 Soares mous
4	378	17.7	430	21	W53498	me81106.r1 Soares mous
5	313	14.6	478	24	AA250120	mx21908.r1 Soares mous
6	303	14.2	515	7	AA512935	mh91b04.r1 NCI_CGAP_br
7	220	10.3	439	12	AA727460	vn96d10.r1 Soares mous
8	160	7.5	459	24	AA245689	mx03607.r1 Soares mous
9	149	7.0	330	9	AA573164	nm51106.r1 NCI_CGAP_br
10	141	6.6	266	7	AA471183	pmv2139 KGLa Lambda za
11	119	5.6	521	18	HI2981	y170a10.r1 Homo sapien
12	100	4.7	229	3	R02249	Y683h05.r1 Homo sapien
13	62	2.9	252	13	AA754459	97SN1787 Rice Immature

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
14	60	2.8	252	13	AA754459	97SN1787 Rice Immature
15	54	2.5	543	24	AA270188	v110602.r1 Soares mous
16	52	2.4	247	13	AA754458	97SN1784 Rice Immature
17	49	2.3	247	13	AA754458	97SN1784 Rice Immature
18	45	2.1	180	23	AA064450	ml47h11.r1 Stratagene
19	36	1.7	284	25	FR0012397	F.rubripes GSS sequenc
20	36	1.7	2275	12	AF034173	Homo sapiens ntcon2 co
21	35	1.6	350	4	AA349976	EST56965 Infant brain
22	29	1.4	478	22	W73681	zds5h01.r1 Soares feta
23	29	1.4	569	25	FR0020899	F.rubripes GSS sequenc
24	24	1.4	596	25	FR0022657	F.rubripes GSS sequenc
25	31	1.4	2275	12	AF034173	Homo sapiens ntcon2 co
26	27	1.3	369	19	N23527	Yv91f01.r1 Homo sapien
27	27	1.3	370	25	AF046247	Mus musculus clone OST
28	27	1.3	416	18	W78296	EST00444 Homo sapiens
29	28	1.3	466	13	AA178124	z146g09.r1 Soares feta
30	27	1.3	553	13	AA568016	HL02152.Sprime HL Dicos
31	26	1.2	193	25	FR0012700	F.rubripes GSS sequenc
32	26	1.2	202	25	FR0014400	EST00010 Tf cDNA libra
33	26	1.2	231	5	W78691	MAAD0155.M3R Schistos
34	26	1.2	301	23	AA218487	MAAD0158.M3R Schistos
35	26	1.2	304	23	AA218491	F.rubripes GSS sequenc
36	26	1.2	304	23	FR0015974	F.rubripes GSS sequenc
37	25	1.2	326	26	HUW078005B	Human STS UT8005, 3' p
38	26	1.2	344	26	HS299YE1	H. sapiens (D185476) D
39	26	1.2	373	7	AA525583	MAAD0146.BAR Schistos
40	26	1.2	382	6	HUM355C04B	Human borta cDNA 5'-en
41	26	1.2	420	13	AA154253	97K00074 Rice Immature
42	26	1.2	544	24	C22979	Dictyostellium discoides
43	26	1.2	547	12	W43683	23300 CD4-16 Arabidops
44	26	1.2	668	5	W78683	EST00013 Tf cDNA libra
45	26	1.2	1287	13	AF038250	Homo sapiens clone ntc

## ALIGNMENTS

RESULT LOCUS	1	AA518362	631 bp	MRNA	EST	16-JUL-1997
DEFINITION		v110602.r1 Barstead mouse myotubes MFLNB5				Mus musculus cDNA clone
ACCESSION		903363 5', MRNA sequence.				
NID		AA518362				
KEYWORDS		92259047				
SOURCE		EST.				
ORGANISM		house mouse.				
REFERENCE		Mus musculus				
AUTHORS		Eukaryotae: mitochondrial eukaryotes; Metazoa: Chordata; Vertebrata: Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae: Mus.				
TITLE		1 (bases 1 to 631)				
JOURNAL		Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucada,T., Lacy,M., Le,M., Martin,D., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
COMMENT		The WashU-HMI Mouse EST Project				
		Unpublished (1996)				
CONTACT:		Marra M/Mouse EST Project				
		WashU-HMI Mouse EST Project				
		Washington University School of Medicine				
		4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
		Tel: 314 286 1800				
		Fax: 314 286 1810				
		Email: mouseest@wustl.edu				
		This clone is available royalty-free through LINT; contact the				
		IMAGE Consortium (info@image.llnl.gov) for further information.				
		MG1:524027				
		Seq primer: -28m13 rev2 ET from Amersham				
		High quality sequence stop: 259.				
FEATURES		Location/Qualifiers				
		1..631				
		/organism="Mus musculus"				
		/strain="C3H"				

```

/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoRI. Site_2: NotI. 1st strand cDNA
was primed with a Not I - 0190(5T) primer [5'
TGTACCAATCTGAAAGGAGCGCGCCCTTTTCTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptor
[AAATCGATCCCTTG], digested with Not I and cloned into
Not I and Eco RI sites of the modified pRT3 vector.
Library constructed by Bob Barstead. The C2C12 cell
(available from ATCC, catalog # CRL-1772) differentiat
rapidly, forming contractile myotubes and producing
characteristic muscle proteins."
/db_xref="taxon:10090"
/clone="903363"
/clone_id="Barstead mouse myotubes MPLRB5"
/cell_line="C2C12"
/lab_host="DH10B"
<1. >631

```

Db	14	GTGCACCAAAAGCCCTCAGGACGATCTTCTTGACAAAGTACACAGCCAAAGCACACACTACGGGAT	73
Oy	874	GTGCACCAAAAGGCCCTCAGGACGATCTTCTTGACAAAGTTCGGGCCAAAGCACACACTACGGGAT	933
Db	74	GCTCTTCTGCTCTCTCCGGGAGCTGCTGCTGCACCGAGAGGCGGAGACAGACTATCTGCC	133
Oy	934	GCTCTTCTGCTCTCTCCGGGAGCTGCTGCTGCACCGAGAGGCGGAGAGACTATCTGCC	993
Db	134	TGTGTGCTCTATGAAAGAGAGAGAGGCCCAACAGCTGTAATCTGCAGAGCTCTTGCA	193
Oy	994	CGTGTGCTCTATGAAAGAGAGAGGCCCAACAGCTGTAATCTGCAGAGCTCTTGCA	1053
Db	194	GACAAATATACATCTGCAGATCTCGCTTGACAGAGTTTTTTTACCAACTGCCAGCAGAGT	253
Oy	1054	GACAAATATACATCTGCAGATCTCGCTTGACAGAGTTTTTTTACCAACTGCCAGCAGAGT	1112
Db	254	CAAGTCTGTACGACAACTGTCTTAAGGAGAACTACGACAGATGCTCTGCTGCTTACCTGG	313
Oy	1113	CAAGTCTGTACGACAACTGTCTTAAGGAGAACTACGACAGATGCTCTCTGCTGCTTACCTGG	1172
Db	314	GACTGATTTGGCAGACATGACTCCTTACTACTATGACTCCAGCAGCCTCAGTGTGGGCG	373
Oy	1173	GACTGATTTGGCAGACATGACTCCTTACTACTATGACTCCAGCAGCCTCAGTGTGGGCG	1232
Db	374	CGTGTGCGATTTGCAGCAGAGTGGCAATGACTGGAAAGATGGCTGTGAATTTCTGAAT	433
Oy	1233	CATGCTGTGACTGCGACGACAAAGCGGCATGACTGGAAAGATGGCTGTGAATTTCTGAAT	1292
Db	434	TTTTTAAAGGCAATACGTGTCTCAAAAATGCAATTTCAAGCCTTTGGCAATGGCTTCGGATG	493
Oy	1293	TTTTTAAAGGCAATACGTGTCTCAAAAATGCAATTTCAAGCCTTTGGCAATGGCTTCAGATG	1352
Db	494	TGACCATGTGGCAGCAGCACACACAC-AGTCAGACACCAATGGCCAGCATCTACATGGCT	552
Oy	1353	TGACCATGTGGCAGCAGCACACACAC-AGTCAGACACCAATGGCCAGCATCTACATGGCT	1412
Db	553	TCGGGATCAGAGCAGACTACTGGG-CCGCGAGGTTCTGAGAAATGAGAT	600
Oy	1413	TCGGGATCAGAGCAGACTACTGGG-CCGCGAGGTTCTGAGAAATGAGAT	1461
RESULT	2		
LOCUS	AA387098	446 bp	mRNA
DEFINITION	vc18b07.t1	Ko mouse embryo 11 5dpc	Mus musculus cDNA clone 774901
ACCESSION	AA387098		
NID	g2040052		
KEYWORDS	EST.		
SOURCE	house mouse.		

ORGANISM	REFERENCE AUTHORS	TITLE JOURNAL	COMMENT
Mus musculus			
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;			
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;			
Murinae; Mus.			
1 (bases 1 to 446)			
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,			
Getsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,			
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,			
Theising, B., Wille, T., Lennon, G., Soares, B., Wilson, R. and			
Waterston, R.			
The Mashu-HIMI Mouse EST Project			
Unpublished (1996)			

Contact: Maria W/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LNC<sup>+</sup>; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:467757  
High quality sequence stop: 413.

FEATURES	Location/Qualifiers
source	1. .446

					/organism="Mus musculus"
					/strain="C57BL/6J"
					/note="Vector: pSPOR1; Site.1: SalI; Site.2: NotI; Total
					RNA were extracted from 11.5 dpc embryos (excluding
					placenta and yolk sac). The double-stranded cDNA was
					synthesized with an oligo (dt)-1 primer
					GAGGAGACCTAGTCTTAGATCGGAGGGCCGCTTTTTTTTTT 3'.
					The cDNAs were ligated to L(-Sal3A: 5'
					GCTATTGACGTGCAGTATTC 3' and L(-Sal3B: 5'
					GGAATGACGACGTCAAT 3'. The cDNAs were size-selected and
					amplified by long-range PCR using Ex Taq polymerase for 18
					cycles. The PCR-amplifiable cDNA mixture went through
					one round of equalization and was digested with SalI/NotI
					and cloned into the SalI/NotI sites of the pSPOR1
					plasmid vector (Life technologies). The library was
					constructed by Dr. Minoru S. H. Ko and Dr. Xiaohong
					Wang."
				/db_xref="taxon:10090"	
				/clone="774901"	
				/clone_1lb="Ko mouse embryo 11 5dpc"	
				/sex="pooled"	
				/dev_stage="11.5dpc"	
				/lab_host="DH10B"	
				<1..>446	
BASE COUNT	112 a	122 c	107 g	105 t	
ORIGIN					mRNA
					Query Match
					Best Local Similarity 95.7%; Pred. NO. 0.00e+00;
					Matches 427; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Db	1	GCAGTCTTCGACAAGTTCACGCCAACACACAGCTACGGGATGCTCTTCTCTCTCGCG	60		19.1%; Score 408; DB 5; Length 446;
Oy	892	GCAGTCTTCGACAAGTTCGCGGCCMACACAGCATCGAGTGCTCTCTCTCGCG	951		
Db	61	GGAGCTCGCTGCACCAGAGGGCGGCGACACTATGCTTCCCTGTGCTCTCTATGAAGA	120		
Oy	952	GGACATCGCTGCACCGGCGGCGGCGACACACTATGCTCCCCGTGTCTCTATGAAGA	1011		
Db	121	ACGAGAGAGGCCCAACTGCTGAACTGTGCAAGACTCTCTGCAAGCAAAATTACATCTGAG	180		
Oy	1012	ACGAGAGAGGCCCAACTGCTGAGCTGCTCAAGAGCTCTGCAAGCAAAATTACATCTGAG	1071		
Db	181	ATCTGCGCTTCGAGATTTTATTACCAACTGGCACGCGAGTCAAGGCTGTGCACAACTG	240		
Oy	1072	ATCTGCGCTTCGAGATTTTATTACCAACTGGCACGCGAGTCAAGGCTGTGCACAACTG	1131		

Db 241 TCTTAGGAGAGACTAGCAGAGACTGCTCTGCTACTCGGAGCTAGTGGCAGATCAT 300  
Oy 1132 TCTTAGGAGAGACTAGCAGAGACTGCTCTGCTACTCGGAGCTAGTGGCAGATCAT 1191  
Db 301 GACTCTCTACTACATAGACTCCAGACGCTCAGTGTGGCGCGGTGGGATTCAGACAA 360  
Oy 1192 GACTCTCTACTACATAGACTCCAGACGCTCAGTGTGGCGCGGTGGGATTCAGACAA 1251  
Db 361 CAGTGGCAGTGGCGGAGAGTGGCTGAGTGTGATTTTAAAGGACAAATACGG 420  
Oy 1232 CAGTGGCAGTGGCGGAGAGTGGCTGAGTGTGATTTTAAAGGACAAATACGG 1311  
Db 421 TCTCAAAATGCAATTCAGACCTTTG 446  
Oy 1312 TCTCAAAATGCAATTCAGACCTTTG 1337  
RESULT 3  
LOCUS W53498 465 bp mRNA EST 03-JUN-1996  
DEFINITION m45e01.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA  
ACCESSION W53498  
KEYWORDS g1357323  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
Murinae; Mus.  
1 (bases 1 to 465)  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
TITLE  
JOURNAL  
COMMENT  
Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:232784  
Seq primer: mob. REGA+ET  
High quality sequence stop: 347.  
Location/Qualifiers  
1. 465  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'-  
TGTACCAATCTGAGTGGAGCGCCGCGGAAATTTTGTGTGTGTGT  
T 3'] , on equal amounts of mRNA from 2 13.5dpc and 2  
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne  
State Univ., from 2.] double-stranded cDNA was ligated to  
Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
pT73 vector. Library went through one round of  
normalization, and was constructed by Bento Soares and  
M.Patima Bonaldo."  
/db\_xref="taxon:10090"  
/clone="371352"  
/clone\_lib="Soares mouse embryo NBME13.5 14.5"  
/sex="unknown"  
/tissue\_type="embryo"  
/dev\_stage="13.5-14.5dpc total fetus"  
/lab\_host="DH10B"

Query Match 18.8%; Score 401; DB 21; Length 465;  
Best Local Similarity 94.0%; Pred. No. 0.0de+00;  
Matches 439; Conservative 0; Mismatches 26; Indels 2; Gaps 2;  
ORIGIN  
mRNA  
BASE COUNT 120 a 133 c 103 g 109 t  
Db 1 GCGTGAATCTGCAAGACTCCGCAAGCAATTTACATCTGAGATCTCGCTGAGATT 60  
Oy 1029 GCGTGAATCTGCAAGACTCCGCAAGCAATTTACATCTGAGATCTCGCTGAGATT 1088  
Db 61 TTTTACCAACTGCGCAGCAGAGTCAAGTCTGTACAGCACTGTCTTAAAGAGAACTACG 120  
Oy 1089 TTTTACCAACTGCGCAGCAGAGTCAAGTCTGTACAGCACTGTCTTAAAGAGAACTACG 1148  
Db 121 CAGACTGCTCTCTG-CCTACTCTGGGAGCTGATGGCAGCATGACTCTTAACTACATAG 179  
Oy 1149 CAGACTGCTCTCTGCGCTACTCTGGGAGCTGATGGCAGCATGACTCTTAACTACATAG 1208  
Db 180 ACTCCAGAGCCTCAGTGTGGCGCGCGTGTGCGATGTGAGCAAGAGTGGCAATGACCTGG 239  
Oy 1209 ACTCCAGAGCCTCAGTGTGGCGCGCGTGTGCGATGTGAGCAAGAGTGGCAATGACCTGG 1268  
Db 240 AAGATTCGCTGAAGTCTGATTTTAAAGCAATACGCTGTCTCAAAATGCAATTC 299  
Oy 1269 AAGATTCGCTGAAGTCTGATTTTAAAGCAATACGCTGTCTCAAAATGCAATTC 1328  
Db 300 AAGCTTTGACATGCTGCGATGTGACCATGTGGCAGCCAGCCCCAGTCCAGACCA 359  
Oy 1329 AAGCTTTGACATGCTGCGATGTGACCATGTGGCAGCCAGCCCCAGTCCAGACCA 1388  
Db 360 CCACATGCGACACTACACTGCTCTCGGATCAAGAAACCTCTAGCGCAGAGGCT 419  
Oy 1389 CCACATGCGACACTACACTGCTCTCGGATCAAGAAACCTCTAGCGCAGAGGCT 1448  
Db 420 CTGAGATGAGATTC-ACACAGCTTTTACACCGTGTGTAATTTG 465  
Oy 1449 CTGAGATGAGATTC-ACACAGCTTTTACACCGTGTGTAATTTG 1495  
RESULT 4  
LOCUS W76716 430 bp mRNA EST 20-JUN-1996  
DEFINITION me81h06.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA  
ACCESSION W76716  
KEYWORDS g1387762  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
Murinae; Mus.  
1 (bases 1 to 430)  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
TITLE  
JOURNAL  
COMMENT  
Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:245779  
Seq primer: mob. REGA+ET

High quality sequence stop: 343.					
FEATURES	Location/Qualifiers				
source	1..430	/organism="Mus musculus"			
		/strain="C57BL/6J"			
		/note="Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15'			
		TGTATCCCATCTGAAAGTGGAAGCGGCCGGCAAAATTTT			
		T 3' ], on equal amounts of mRNA from 2 13.5dp			
		14.5dp embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2 ] double-stranded cDNA was ligated			
		Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified			
		pRTT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."			
		/db_xref="taxon:10090"			
		/clone="402011"			
		/clone_1ib="Soares mouse embryo NbMEI3.5 14.5"			
		/sex="unknown"			
		/tissue_type="embryo"			
		/dev_stage="13.5-14.5dp total fetus"			
		/lab_host="DH10B"			
		<1..->430			
BASE COUNT	112 a 121 c 100 g 97 t				
ORIGIN					
Query Match	17.7%:	Score 378:	DB 21:	Length 430:	
Best Local Similarity	94.4%;	Pred. No. 0.00e+00:			
Matches	407;	Conservative 0;	Mismatches 23;	Indels 1;	Gaps 1;
Db	1 GCCTGAATCTGCAGACTCTCTCAGAACAATAATACATCTGCAGATCTGCCTTCAGAGATT	60			
Oy	1029 GCGTAGTCTGCAAGACTCTCTCAGAACAATAATACATCTGCAGATCTGCCTTCAGAGATT	1088			
Db	61 TTTTACCACACTGCAGCCAGCAGTAAGGCTGTGAGCAACTGTCTTAAAGAGAACTAGC	120			
Oy	1089 TTTTACCACAACTGCAGCCAGCAGTAAGGCTGTGAGCAACTGTCTTAAAGAGAACTAGC	1148			
Db	121 CAGACTGGCTCTGGCCATCTGGGACATGATTGGCAAGTATGACTCTACTACTACATAG	180			
Oy	1149 CAGACTGGCTCTGGCCATCTGGGACATGATTGGCAAGTATGACTCTACTACTACATAG	1208			
Db	181 ACTCCAGACAGCCTCAGTGTGAGCGCGGTGGATTCAGCAACAAGTGGACATGACTGG	240			
Oy	1209 ACTCCAGACAGCCTCAGTGTGAGCGCGGTGGATTCAGCAACAAGTGGACATGACTGG	1268			
	241 AAGATTGCCGTGAAGTTCTGAATTTTTTAAAGA-AATACGTGTCTCAAAAATGCAATTC	299			
Oy	1269 AAGACTGCTTGAATTTCTGAATTTTTTAAAGAACAAATCTGTCTCAAAAATGCAATTC	1328			
Db	300 AAGCTTTTGCAGATGGCTTCGATGTGACATCTGGCAGCACCCCCCAAGTCCAGACCA	359			
Oy	1329 AAGCTTTTGCAGATGGCTTCGATGTGACATCTGGCAGCACCCCCCAAGTCCAGACCA	1388			
Db	360 CCACAGCCACGACTACACTGCTTCGCGATCAAGAACAAAGCTCTAGGGCCAGAGAGCT	419			
Oy	1389 CCACAGCCACGACTACACTGCTTCGCGATCAAGAACAAAGCTCTAGGGCCAGAGAGCT	1448			
Db	420 CTGAGAATGAG 430				
Oy	1449 CTGAGAATGAG 1459				
RESULT	5	AA250120	478 bp	mRNA	EST 12-MAR-1997
LOCUS	mx21g08.tl	Soares mouse NML Mus musculus cDNA clone 680894 5'			mRNA
DEFINITION		sequence.			
ACCESSION		AA250120			
NID		g1882418			
KEYWORDS		EST.			
SOURCE		house mouse.			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 478)	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Eukaryotes: mitochondria eukaryotes: Metazoa: Chordata: Vertebrata: Euthera: Rodentia: Sciurognathi: Muridae: Murinae; Mus.		
2	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.			
3	The Mashu-HMI Mouse EST Project			
4	Unpublished (1996)			
5	Contact: Marra M/Mouse EST Project			
6	Mashu-HMI Mouse EST Project			
7	Washington University School of Medicine			
8	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			
9	Tel: 314 286 1800			
10	Fax: 314 286 1810			
11	Email: mouseest@wustl.edu			
12	This clone is available royalty-free through LINDA; contact the			
13	IMAGE Consortium (info@image.llnl.gov) for further information.			
14	MG1:420598			
15	Seq primer: -28ml3 rev2 ET from Amersham			
16	High quality sequence stop: 470.			
17	Location/Qualifiers			
18	1. 478			
19	/organism="Mus musculus"			
20	/note="Vector: pT73d-Pac (Pharmacia) with a modified			
21	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA			
22	was primed with a Not I - oligo(dT) primer [5',			
23	TGTACCAATCTGAAAGTGGAGCGCGCCGACATCTTTTCTTTTCTT 3'];			
24	double-stranded cDNA was ligated to Eco RI adaptors			
25	(Pharmacia), digested with Not I and cloned into the Not I			
26	and Eco RI sites of the modified pT73d vector. Library			
27	constructed and normalized by Bento Soares and M.Fatima			
28	Bonaldo."			
29	/db_xref="taxon:10090"			
30	/clone="680894"			
31	/clone_lib="Soares mouse NML"			
32	/tissue_type="Liver"			
33	/lab_host="DH10B"			
34	<1. ->478			
35	BASE COUNT	121 a	116 c	111 g
36	ORIGIN			130 t
37	Query Match	14.6%	Score 313;	DB 24; Length 478;
38	Best Local Similarity	89.3%	Prod. No. 0.00e+00;	
39	Matches 424; Conservative	0;	Mismatches 39;	Indels 12; Gaps 4;
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				

Db	353	TAAGAAACCTTTTTCGCGCCCTCACAAGGCTTGTTGAAGAACCTGCTACAGGCTAAATTC	412
Oy	1846	TAAGAAACCTTGT-G-GCCCTCAGAGGCTTGTTGTAAGAACTCTACAGGCTAAATTC	1902
Db	413	CAAACTCAGAAACCTCTGGGCGCATGGTGTGCGCTTACAGGAGACTGTTGTAACTCG	467
Oy	1903	CAAAACCATTAAAGCGCTCTGGGCGCTGTGCGGCTTAAGGGAGACATTTCACCATG	1957
RESULT	6		
LOCUS	AA512935	515 bp	EST
DEFINITION	nh11b04.s1 NCI_CGAP_Brl.1 Homo sapiens CDNA clone IMAGE:965839,		05-AUG-1997
	MRNA sequence.		
ACCESSION	AA512935		
NCID	92251358		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;		
	Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;		
	Homo.		
REFERENCE	1 (bases 1 to 515)		
AUTHORS	NCI-CGAP hhttp://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT			

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: [Robert\\_Strausberg@nih.gov](mailto:Robert_Strausberg@nih.gov)  
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CCG clone distribution information can be  
found through the I.M.A.G.E. Consortium/TLN! at:  
[www.bio.linn.gov/bdrp/image/image.html](http://www.bio.linn.gov/bdrp/image/image.html)

Insert Length: 1735	Std Error: 0.00
Seq primer: ~40m3 fwd. ET from Amerham	
High quality sequence stop: 460.	
Location/Qualifiers	
source	1. .515

```

/organism="Homo sapiens"
/note="Vector: pRTT3-Pac (Pharmacia) with a modified
polylinker: 1st strand cDNA was prepared from pooled bulk
breast tumor tissue, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pRTT3
vector. Library is not normalized. (The normalized
version of this library is NCI CGAP Br2.) Library was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_image:965839"
/clone_11b="NCI CGAP Br1.1"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
<1. >515

```

Query Match 14.2%; Score 303; DB 7; Length 515;  
 Best Local Similarity 85.0%; Pred. No. 0.00e+00;  
 Matches 398; Conservative 0; Mismatches 65; Indels 5; Gaps 3;

[illegible]

RESULT	7	AA727460	439 bp	mRNA	EST	02-JAN-1998
LOCUS		v966d10.r1		Stratagene mouse skin (#937313)	Mus musculus	CDNA clone
DEFINITION		1210003.5		mRNA sequence.		
ACCESSION		AA727460				
NID		92745167				
KEYWORDS		EST.				
SOURCE		house mouse.				
ORGANISM		Mus musculus				

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
1 (bases 1 to 439) Marra, M., Hallier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Gelsel, S., Kucaba, T., Lacy, M., Martin, J., Morris, M., Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B., Thaistang, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Wrepton, R.	The WashU-HMII Mouse EST Project Unpublished (1996)

Contact: Marra M/Mouse EST Project  
 WashU-HHMI Mouse EST Project  
 Washington University School of Medicinep  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel.: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LINT ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:646347  
 .Seg primer: -28n13 rev1 ET from Amersham  
 High quality sequence stop: 436.  
 Location/Qualifiers  
 1. .439

```

/note="Organ: skin; Vector: plasmidscript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: 01590 dt. Whole skin from 11 week old C57BL/6 female mice. Average insert size: 1.0 kb; In-vap XR Vector: -5' adaptor sequence: 5' GAATTCGCACGCG 3' -3' adaptor sequence: 5' CTCGACGTTTTTTTTTTTTTTTTT 3."
/db_xref="taxon:10090"
/clone="1210003"

```



[illegible]

**FEATURES**

<b>Source</b>	<b>location/Qualifiers</b>		
	1..459		
	/organism="Mus musculus"		
	/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5], TGTTACCAATCTGAGCGAGGGCCGCCAATTCTTTTTTTTTTTT 3'; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Library constructed and normalized by Bento Soares and M.Palmira Bonaldo."		
	/db_xref="taxon:10090"		
	/clone="679140"		
	/clone_1id="Soares mouse NML"		
	/tissue_type="Liver"		
	/lab_host="DH10B"		
	<1..>	459	

BASE COUNT	117 a	87 C	118 G	137 T
ORIGIN				
Query Match	7.5%; Score 160; DB 24; Length 459;			
Best Local Similarity	80.1%; Pred. No. 6,376-300;			
Matches	226; Conservative	0; Mismatches	54; Indels	2; Gaps 1;

```

Db    13 TTGC GGCCCTCACA AGGCTTCGTGTGAAGA CTGCTACAGSGCTAATTCCAA CTCAGAAA 72
      |||         |             |            |           |          |||
Qy   1855 TTGG CGCCCTCAGG GGGCTTC GTTGAA GAAGACTGCTACAGSGCTAATTCCAA ACCCATPAG 1914

Db    73 ACTGTGGGC ATGTGTGTCGTTCAGGGGACGTGTTGTTA CCTGGAC TCYTAAGCAAGCTG 132
      |||||       |     |||||||  |||||        |||  |||  |||  |||  |||  |||
Qy   1915 GCCTGGGG CCCTGGTGC GGCTTAAAGGGACCATTG GCCACTGTGA AACCAAGCTGGGCTT 1974

Db    133 AGCTTATGTTTGTATGATGTGATGCTGCGTAGATGATAAGATATTTTA ACAGCTT 192
      ||  |||      |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy   1975 ATC--ATGTGTTGATGTG TAGAGATGCTAGTGATGCTATGTATTTT TAACAGCTT 2032

Db    193 GAACCTGTGTCCTCTACTAGTGGCTAGAGAGAGAGAGATGCTATTGATAGATTCCTTC CGGA 252
      |||||       |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||
Qy   2033 GAACCTGTGTCCTCTACTAGTGGCTAGAGAGAGAGATGCTATTGATTAAGATTCCTTC CATG 2092

Db    253 TCTTATTTTAATAGCGTTCGATTCTTAAAGACATGACTGACGCC 294
      |||||       |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy   2093 TCTTACTCAGCACGATTCCTTCGAAAGACAGGCCCGCAGCC 2134
  
```

RESULT LOCUS DEFINITION ACCESSION MID KEYWORDS SOURCE ORGANISM	9 AA573164 mm51106.s1 sequence.	330 bp mRNA Homo sapiens	EST	28-AUG-1997
REFERENCE AUTHORS TITLE JOURNAL COMMENT	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)	Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmet-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be		









\*\*\*\*\*  
[W27327] (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSEARCH protein - protein database search, using Smith-Waterman algorithm  
on: Tue Jun 23 18:34:40 1998: Maspar time 15.71 Seconds  
456.813 Million cell updates/sec  
Molecular output not generated.

Title: >US-08-866-354-4  
Description: (1-468) from US08866354.pep  
Perfect Score: 3384  
Sequence: 1 MFLLATLYFALPLDLMSAE.....PVLMATLALLSVLAETS 468

Scoring table:  
Gap 11  
PAM 150

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 08  
Listing first 45 summaries

Database: a-geneseq31.2  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 34.847; Variance 146.333; scale 0.238

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	3384	100.0	468 27	W27327	Rat glial cell derive	0.00e+00
2	95	2.8	350 15	R82686	Vascular endothelial	2.23e+01
3	96	2.8	505 1	P80885	Sequence encoded by L	1.89e+01
4	91	2.7	316 24	W15023	C-mpl ligand variant	4.31e+01
5	90	2.7	332 20	W12929	Mpl ligand analogue,	5.06e+01
6	92	2.7	369 3	P50258	Pseudorabies virus (B	3.66e+01
7	90	2.7	277 3	W35853	Human CD18 for use in	5.06e+01
8	90	2.7	769 4	R24256	Beta-subunit CD18 of	5.06e+01
9	91	2.7	947 27	W42402	Amino acid sequence o	4.31e+01
10	87	2.6	32 3	R15252	Carbohydrate binding	8.20e+01
11	88	2.6	95 18	W00601	Rat carboxypeptidase	6.99e+01
12	87	2.6	285 21	W01770	IL-3 variant/Fxa/clase	8.20e+01
13	87	2.6	285 21	W01771	IL-3 variant/GlySer/1	8.20e+01
14	87	2.6	332 21	W01760	Variant c-mpl ligand	8.20e+01
15	87	2.6	504 3	P93707	Sequence of the gag p	8.20e+01
16	87	2.6	514 3	R15237	Fusarium oxysporum C	8.20e+01
17	88	2.6	321 1	P81769	Sequence encoded by t-	6.99e+01
18	86	2.5	119 27	W43305	p17 antigen of human	9.61e+01
19	86	2.5	120 2	P70272	The sequence encoding	9.61e+01

20	86	2.5	132 5	R26326	HTLV-IIIB p17 sequenc	9.61e+01
21	86	2.5	169 10	R55855	HIV-1 MA delta-79-90	9.61e+01
22	86	2.5	169 10	R55851	HIV-1 MA delta-5-16 m	9.61e+01
23	86	2.5	171 10	R55857	HIV-1 MA delta-105-11	9.61e+01
24	86	2.5	267 2	P94509	Sequence encoded by p	9.61e+01
25	86	2.5	332 20	W12942	Mpl ligand analogue,	9.61e+01
26	86	2.5	363 8	R43267	p37gag encoded by vec	9.61e+01
27	86	2.5	438 13	R71783	HIV-1 gag protein p45	9.61e+01
28	86	2.5	478 13	R73878	HIV-1 gag polyprotein	9.61e+01
29	86	2.5	478 3	P60309	Sequence of gag and p	9.61e+01
30	86	2.5	500 2	P70269	The sequence encoding	9.61e+01
31	86	2.5	506 6	R29704	gag gene decoded from	9.61e+01
32	86	2.5	506 3	P60176	Sequence of ARV-2 (98	9.61e+01
33	86	2.5	508 3	P60419	Sequence of LAV virus	9.61e+01
34	86	2.5	512 20	W02296	HIV-gag internal prot	9.61e+01
35	86	2.5	512 4	R24028	HIV-gag protein.	9.61e+01
36	86	2.5	512 2	R10866	HIV gag precursor pro	9.61e+01
37	86	2.5	515 1	P91235	(ENV-80)(GAG-VII)(Hex	9.61e+01
38	86	2.5	522 2	R04024	Gag gene product of c	9.61e+01
39	86	2.5	524 27	W43304	Modified p17 antigen	9.61e+01
40	86	2.5	600 2	P70541	HTLV-III gag/env gene	9.61e+01
41	86	2.5	607 7	R34662	Rat 5HT transporter.	9.61e+01
42	86	2.5	639 1	P91052	Sequence of fusion pr	9.61e+01
43	86	2.5	769 1	P80836	Beta subunit of human	9.61e+01
44	86	2.5	769 1	R07113	Recombinant beta-subu	9.61e+01
45	86	2.5	3077 1	P93283	Sequence of clone HIV	9.61e+01

## ALIGNMENTS

RESULT	ID	W27327	standard; Protein; 468 AA.
AC	W27327	1	
DE	27-MAR-1998	(first entry)	
DE	Rat glial cell derived neurotrophic factor receptor alpha.		
KW	Glial cell derived neurotrophic factor receptor alpha; GDNF alpha;		
KW	GDNF; rat; kidney disease; glomerulonephritis; therapy.		
OS	Rattus sp.		
PH	Key	Location/Qualifiers	
FT	Peptide	1..24	
FT	Domain	/label= Sig_peptide	
FT	Domain	25..445	
FT	Domain	/note= "extracellular domain"	
FT	Domain	446..468	
FT	Domain	/note= "GPI-attachment site"	
FT	Modified_site	61	
FT	Modified_site	/note= "O-glycosylated"	
FT	Modified_site	349	
FT	Modified_site	/note= "O-glycosylated"	
FT	Modified_site	408	
FT	Modified_site	/note= "O-glycosylated"	
FT	Cleavage_site	428..430	
FT	Cleavage_site	/note= "GPI-anchor cleavage/attachment site"	
PN	W09733912-A2.		
PD	18-SEP-1997.		
PF	13-MAR-1997: U04363.		
PR	14-MAR-1996; US-618236.		
PA	14-MAR-1996; US-615902.		
PR	(GERTH ) GENENTECH INC.		
PI	Klein RD, Moore MW, Rosenthal A, Ryan AM;		
PI	WPI: 97-470819/43.		
DR	N-PSDB: T84975.		
PT	Isolated glial cell derived neurotrophic factor receptor alpha -		
PT	useful to develop products to diagnose and treat associated		
PT	disorders, particularly enteric nervous system or kidney disorders		
PS	Claim 1: Page 78-79; 100pp; English.		
CC	This protein comprises full-length rat glial cell derived		
CC	neurotrophic factor receptor alpha (GDNF alpha), a novel		
CC	GPI-linked protein that is a ligand-binding component of the		
CC	receptor system for GDNF. Its amino acid sequence was deduced		
CC	from an isolated cDNA clone (see T84975). The invention relates		
CC	to novel uses of GDNF and its receptor. In particular, it		
CC	relates to native rat GDNF alpha (see W27327), its variants and		

CC soluble derivatives (extracellular domain), chimeric GDNF alpha  
 CC and antibodies which bind to the GDNF alpha, including agonist  
 CC and neutralising antibodies, as well as various uses for these  
 CC molecules. It also relates to assay systems for detecting ligands  
 CC to GDNF alpha, systems for studying the physiological role of  
 CC GDNF, diagnostic techniques for identifying GDNF-related conditions,  
 CC methods for identifying molecules homologous to GDNF alpha, and  
 CC therapeutic techniques (claimed) for the treatment of GDNF-related  
 CC and GDNF alpha-related conditions, particularly kidney disease  
 CC associated with glomerulonephritis and enteric nervous system  
 CC related disorders. Transgenic and knockout animals are also  
 CC claimed.

CC Sequence 468 AA:

Query Match 100.0%: Score 3384; DB 27; Length 468;  
 Best Local Similarity 100.0%: Pred. No. 0.00e+00;  
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 mfacllyfaipldlmsaevsgdrldcvkaadqclkegscstkrrtlrvcagketnf 60  
 1 MFLATLYFALPLDLMLSAEVSQGGDRIDCVKASQCLKEGSCSTKRRTLRQCVAGKETNF 60  
 DB 61 sltvglaekdecraamealqkalsyncrckrmkknclrlywmyqslqgnldisp 120  
 61 SLTSGLEAKDECRAAMEALQKSLYNCRCKRMKKNCLRLIYWSMYQSLQGNLDLSDSP 120  
 QY 121 YEPVNSRLSDIFRVPFISDVFQGVHLSKYNCLDAKACNLDITCKRKSATITPCTT 180  
 121 YEPVNSRLSDIFRVPFISDVFQGVHLSKYNCLDAKACNLDITCKRKSATITPCTT 180  
 DB 181 smsnevevnrckhkalrfdkvpakhsygmllfscrcdiacterrrrgtllypvcyeer 240  
 181 SMSNEVEVNRCKHKAIRFDKVPAKHSYGMLLFSCRCDIACTEERRRRGTIYPVCYEER 240  
 QY 241 PNCISLQDSCNTKYNICRSRLADFEFTNCPESRSYCNCKENYACDLAYSLGTLVTMPN 300  
 241 PNCISLQDSCNTKYNICRSRLADFEFTNCPESRSYCNCKENYACDLAYSLGTLVTMPN 300  
 QY 301 YVDSSSLVAPMCCSSGNDLEDCLEFLNFKDNTCLKNALQAFNGSDVTMMQAPPV 360  
 301 YVDSSSLVAPMCCSSGNDLEDCLEFLNFKDNTCLKNALQAFNGSDVTMMQAPPV 360  
 DB 361 qtttaatttafrvknkpljpagseeneipthvlppcanlgaqklkxvsgsthlclsdgf 420  
 361 OTTATATTAFRVKNKPLJPAGSENEIPTHVLPPCANLGAQKLKXVSGSTHCLSDGF 420  
 QY 421 gkgglagasshltcksmappscslslpvlmltalalaallsvslaets 468  
 421 GKDGLAGASSHITTKSMAPPSCSLSLPVLMLTALALSVSLAETS 468

RESULT 2  
 ID R82686 standard; Protein: 350 AA.

AC R82686;  
 DT 18-APR-1996 (first entry)  
 DE Vascular endothelial growth factor 2.  
 KW VEGF2; vascular endothelial growth factor; vascular permeability factor;  
 KM PDGF family; angiogenesis; induce growth; bone.

OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..24  
 FT /label= signal\_peptide  
 FT protein 25..350  
 FT /label= mature\_protein  
 FT misc\_difference 62  
 FT /note= "conserved cys"  
 FT domain 85..98  
 FT /note= "signature for the PDGF/VEGF family"  
 FT misc\_difference 87  
 FT /note= "conserved cys"  
 FT misc\_difference 93  
 FT /note= "conserved cys"  
 FT misc\_difference 96  
 FT /note= "conserved cys"

FT /note= "conserved cys"  
 FT misc\_difference 97  
 FT /note= "conserved cys"  
 FT misc\_difference 104  
 FT /note= "conserved cys"  
 FT misc\_difference 140  
 FT /note= "conserved cys"  
 FT misc\_difference 142  
 FT /note= "conserved cys"  
 FT /note= "conserved cys"  
 PN W09524473-A1.  
 PD 14-SEP-1995.  
 PR 12-MAY-1994; U05291.  
 PR 08-MAR-1994; U05291.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Cao L, Hu T, Fraser CM;  
 DR WPI: 95-328266/42.  
 DR N-PSDB: T03950.  
 PT Human vascular endothelial growth factor 2 polypeptide - used to  
 PT provide diagnostic, therapeutic and/or prophylactic effects against  
 PT various diseases, esp. in wound healing and periodontal disease  
 PS Claim 1: Fig 1A-D: 48pp; English.  
 CC This full length vascular endothelial growth factor 2 (VEGF2) comprises  
 CC 350 amino acids of which approx. the first 24 amino acids represent the  
 CC leader sequence. VEGF2 is structurally related to the VEGF/PDGF family.  
 CC It is partic. important that all eight cysteines are conserved within  
 CC the members of the family. In addition, the signature for the PDGF/VEGF  
 CC family, PXCXXXXRCGCGCN, is conserved in VEGF2. VEGF2 induces growth of  
 CC damaged bone, peritendium or ligament tissue. It is used for the  
 CC angiogenesis is important in keeping wounds clean and non-infected, VEGF2  
 CC may be used in association with surgery and following the repair of cuts.  
 SQ Sequence 350 AA:

Query Match 2.88; Score 95; DB 15; Length 350;  
 Best Local Similarity 23.28; Pred. No. 2.33e+01;  
 Matches 13; Conservative 14; Mismatches 26; Indels 3; Gaps 3;

DB 8 ywkmkqclrkqgwmnreganlnsrteefkfrhaahnteilksidnewrktqcm 63  
 103 YWKMKQCLRKQGWMNREGANLNSRTEEFKFRHAAHNTEILKSIDNEWRKTQCM 63

RESULT 3  
 ID P80885 standard; Protein: 505 AA.  
 AC P80885;  
 DT 16-DEC-1990 (first entry)  
 DE Sequence encoded by LAV MA L GAG gene  
 KW HIV; HTLV III; AIDS; diagnosis; vaccine; probe; hybridisation.  
 OS Lymphadenopathy associated virus MA L.  
 PN W08707906-A.  
 PD 30-DEC-1987.  
 PF 22-JUN-1987; E00326.  
 PR 23-JUN-1986; EP-401380.  
 PA (INSP) Inst Pasteur.  
 PI Alizon M, Sonigo P, Wain-Hobson S, Montagnier L;  
 DR WPI: 88-014396/02.  
 DR N-PSDB: N80437.  
 PT New variants of lymphadenopathy associated virus (LAV) -  
 PT used for prodn. of DNA, antigens and antibodies used in  
 PT diagnosis of AIDS and pre-AIDS  
 PS Claim 8: Fig 8A-8I: 72pp; English.  
 CC LAV EL I (N80436) and LAV MA L (N80437) were isolated from the peripheral  
 CC blood lymphocytes of patients. Different AIDS virus isolates concerned  
 CC are designated by 3 letters of the patients name. Stable probes including  
 CC the DNA sequences can be used for detection of the new LAV viruses or  
 CC related viruses or DNA proviruses in eg. biological samples. The proteins  
 CC or peptides can be used for detection of antibodies induced in vivo and  
 CC present in biological fluids. The DNA can also be used for the expression  
 CC of LAV viral antigens for the prodn. of a vaccine against LAV. The  
 CC polypeptides can also be used for the prodn. of antibodies for the  
 CC detection of proteins related to the LAV viruses, partic. for diagnosis  
 CC of AIDS or pre-AIDS.  
 SQ Sequence 505 AA:





[illegible][illegible]



Query Match	2.6%	Score 87;	DB 21;	Length 285;
Best Local Similarity	22.0%	Pred. No. 8.20e+01;		
Matches	27;	Conservative	33;	Mismatches 59; Indels 4; Gaps 4;
Db	124	gagggsmasapapacdlrvlsklirshvlsrsgcpevhlpfpvllpavdfsjgew	183	
OY	346	GNQSDVTFWMQAPAPV-QTTTATTT-TAFRVKYNKPLGPAAGSENEIPTHVLPCCANLQAKL	403	
Db	184	ktgmeektagqdlisavlllleqymaarqgqipctcissllg-qlsqgyrlllgaalqslgt	242	
OY	404	KSNVSGS-THLCLSDSPFGKDLGAGASHHTTKSMAAPPCSCSLSLPVLMTLALALLSV	462	
Db	243	qlp 245		
OY	463	SLA 465		
RESULT	13	W01771 standard; Protein: 285 AA.		
DE	12-MAY-1997	(first entry)		
KW	c-mpl ligand; Megakaryocyte Growth and Development Factor; MGF; thrombopoietin; TPO; variant; muten; chimera; increased activity; decreased side effect; ex vivo expansion; stem cell; treatment; haematopoietic disorder; gene therapy; human; interleukin-3.			
OS	Synthetic.			
PN	W06923888-A1.			
PR	08-AUG-1996.			
PR	01-FEB-1996; U00830.			
PA	03-FEB-1995; US-383035.			
PI	(SEAR.) SEARLE & CO G.D.			
PI	Baum CM, Favara JP, Kahn LE, McKearn JP, Pegg LE; Staten NR;			
DR	WPI; 96-371436/37.			
DR	N-PSDB; T59356.			
PT	Mutant c-mpl ligands - used for stimulating the proth. of haematopoietic cells and in the treatment of haematopoietic disorders			
PS	Example 13; Page -: 74pp; English.			
CC	This protein comprises interleukin-3 variant 13288 fused to a synthetic linker and amino acids 1-153 of c-mpl ligand. It is encoded by E. coli expression plasmid, pMON26471. The native c-mpl ligand is also referred to as Megakaryocyte Growth and Development Factor (MGDF) or thrombopoietin (TPO). Variants and chimera of c-mpl ligand can have an improved biological profile, such as increased activity and/or improved side effects, and/or improved physical properties, such as improved half-life, stability and/or re-fold efficiencies. They can be used for selective ex vivo expansion of stem cells, for the treatment of patients having a haematopoietic disorder or in human gene therapy.			
CC	Sequence	285 AA;		
SC	Query Match	2.6%;	Score 87;	DB 21;
	Best Local Similarity	22.0%;	Pred. No. 8.20e+01;	
	Matches	27;	Conservative	33;
			Mismatches 59;	Indels 4;
			Gaps 4;	
Db	124	gagggsmasapapacdlrvlsklirshvlsrsgcpevhlpfpvllpavdfsjgew	183	
OY	346	GNQSDVTFWMQAPAPV-QTTTATTT-TAFRVKYNKPLGPAAGSENEIPTHVLPCCANLQAKL	403	
Db	184	ktgmeektagqdlisavlllleqymaarqgqipctcissllg-qlsqgyrlllgaalqslgt	242	
OY	404	KSNVSGS-THLCLSDSPFGKDLGAGASHHTTKSMAAPPCSCSLSLPVLMTLALALLSV	462	
Db	243	qlp 245		
OY	463	SLA 465		
RESULT	14	W01760 standard; protein: 332 AA.		
ID	W01760;			
AC	W01760;			
DT	23-APR-1997 (first entry)			

```

DE Variant: c-mpl ligand (1-332) P46L+M200R.
KW c-mpl ligand; Megakaryocyte Growth and Development Factor; MGDF;
KW thrombopoietin; TPO; variant; mutant; chimera; increased activity;
KW decreased side effect; ex vivo expansion; stem cell; treatment;
KW haematopoietic disorder; gene therapy; human.
OS Synthetic.
FH Key
FT misc_difference 281.300
FT Location/Qualifiers
FT misc_difference 46
FT /note="substitution of Leucine for wild type residue"
FT /note="encoded by T59343, missing from the protein
FT sequence referred to in the specification"
FT misc_difference 46
FT /note="substitution of Leucine for wild type residue"
PD WC9623888-A1.
PD 08-AUG-1996.
PF 01-FEB-1996; U00830.
PF 03-FEB-1995; US-383035.
PA (SEAR ) SEARLE & CO G D.
PI Baum CM, Favara JP, Kahn LE, McKearn JP, Pegg LE;
PI Stelen NR.
PI WPI: 96-37136/37.
DR N-PSDB: T59343.
PT Mutant c-mpl ligands - used for stimulating the produ. of
PT haematopoietic cells and in the treatment of haematopoietic
PT disorders
PS Example 2: Page -?: 74pp; English.
CC This protein comprises amino acids 1-332 of c-mpl ligand with a P46L
CC (sic) +W400R substitution. It is encoded by BHK expression plasmid
CC PMON2451-4 (T59343). It is a specific example of the variant c-mpl
CC ligands of the invention. The native c-mpl ligand is also referred to
CC as Megakaryocyte Growth and Development Factor (MGDF) or thrombopoietin
CC (TPO). Variants and chimera of c-mpl ligand can have an improved
CC biological profile, such as increased activity and/or decreased side
CC effects, and/or improved physical properties, such as improved half-life
CC stability and/or re-fold efficiencies. They can be used for selective ex
CC vivo expansion of stem cells, for the treatment of patients having a
CC haematopoietic disorder or in human gene therapy.
SQ Sequence 332 AA:

Query Match 2.68; Score 87; DB 21; Length 332;
Best Local Similarity 22.18; Pred. No. 8.20e+01;
Matches 21; Conservative 28; Mismatches 44; Indels 2; Gaps 2;

Db 20 hvhsrlsqcpevpltpylpavdlsgewktmeekagdlgavelllegymaary 79
Oy 372 RYKKNKPLGPAGSENEIRTFHVLPPCANLQAQKLSNVSGS-THLCLSDSPFGKDLGAGSS 430
Db 80 qlptcslsllg-qlsgqvlllgaalgsllgtqlp 113
Oy 431 HITKSMAPRPSCLSSLPVIMLTALALSLVSLA 465

RESULT 15
ID P93707 standard; Protein; 504 AA.
AC P93707;
DC 11-MAR-1992 (first entry)
DE Sequence of the gag protein precursor.
KW Vaccine; diagnosis; AIDS; LAV; HTLV-III.
OS Human immunodeficiency virus.
PN EP-345242-A.
PD 06-DEC-1989.
PF 02-JUN-1989; 870082.
PF 03-JUN-1988; US-202271.
PR 30-JUN-1989; AU-037256.
PR 11-JUL-1989; WO-002415.
PA (SMIK ) SMITHKLINE BIOLOGIC.
PI Cheysen DA, Jacobs E;
PI WPI: 89-358596/49.
DR N-PSDB: T59343.
PT New recombinant DNA encoding HIV gag precursor protein - without
PT flanking sequences, expressed in insect, yeast and mammalian
PT cells, producing particles useful in vaccines and diagnosis
PT Disclosure, Page 4-6; 17pp; French.
CC The inventors claim recombinant gag protein precursor (P93707) and

```



THIS PAGE BLANK (USPTO)

\*\*\*\*\*  
 WISE (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPERCH\_PP protein - protein database search, using Smith-Waterman algorithm  
 on: Tue Jun 23 18:32:21 1998: Maspar time 23.69 Seconds  
 721.562 Million cell updates/sec  
 Modular output not generated.

Title: >US-08-866-354-4  
 Description: (1-468) from US08866354.pep  
 Perfect Score: 3384  
 Sequence: 1 MFLATLYFALPLDLMSAE.....PYMLTRALALSVSLAETS 468

Scoring table:  
 PAM 150  
 Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: p1r56  
 1:p1r1 2:p1r2 3:p1r3 4:p1r4 5:nr13d

Statistics: Mean 46.167; Variance 82.676; scale 0.558

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	178	5.3	24	2	S69080	glycosyl-phosphatidyl
2	108	3.2	397	2	S70987	dnan protein - Mycoba
3	102	3.0	249	2	S09868	hypothetical protein
4	103	3.0	399	2	S70983	dnan protein - Mycoba
5	100	3.0	522	2	S41819	nucleoporin p62 - hum
6	99	2.9	352	2	S60024	bradykinin B1 recepto
7	95	2.8	169	2	I64089	protein-export protei
8	96	2.8	769	1	JC1121	leukocyte adhesion pr
9	96	2.8	826	2	A60385	monocyte surface anti
10	94	2.8	857	2	S33821	median body protein -
11	94	2.8	1186	2	S72329	meiotic recombination
12	92	2.7	100	2	S26728	hypothetical protein
13	91	2.7	316	2	S58719	probable membrane pro
14	90	2.7	327	2	A55356	urokinase-type plasmi
15	92	2.7	336	2	D70030	alkanal monooxygenase
16	91	2.7	346	2	A70144	hypothetical protein
17	91	2.7	469	2	I37451	HBF-G2 (HFK-2) protei
18	90	2.7	475	2	S49886	probable membrane pro
19	90	2.7	493	2	JC5621	epidermal growth fact
20	93	2.7	600	2	S07638	sport coat protein SP
21	90	2.7	600	2	S75664	hypothetical protein
22	92	2.7	688	1	C1HUS	complement subcompone
23	93	2.7	695	2	S05008	serine proteinase, ca

24	91	2.7	772	2	S32659	Integrin beta 2 chain	4.80e+00
25	93	2.7	798	2	S09867	hypothetical protein	6.45e+00
26	90	2.7	1075	2	S54067	probable membrane pro	1.53e+01
27	87	2.6	293	2	R26637	neurogenic repetitive	1.15e+01
28	88	2.6	378	2	S00842	leukostallin precursor	1.15e+01
29	87	2.6	411	2	H69158	LPS biosynthesis Ribu	1.53e+01
30	88	2.6	415	2	A32129	carboxypeptidase B (E	1.15e+01
31	88	2.6	448	2	S57909	probable histidine pr	1.15e+01
32	87	2.6	477	1	R46396	transcription factor	8.63e+00
33	89	2.6	491	2	S49779	CCO1 protein - yeast	8.63e+00
34	88	2.6	506	1	A38068	gag polypeptide - hum	1.15e+01
35	87	2.6	525	2	A35396	nuclear pore glycopro	1.53e+01
36	89	2.6	563	1	VCMVM7	env polypeptide - Bab	8.63e+00
37	88	2.6	743	2	A48917	probable regulatory p	1.15e+01
38	89	2.6	770	2	S04847	leukocyte adhesion pr	8.63e+00
39	89	2.6	771	2	A45839	leukocyte adhesion pr	8.63e+00
40	88	2.6	1188	2	F64367	pyruvate, water dikina	1.15e+01
41	87	2.6	2139	2	A35672	crumbs protein - frui	1.53e+01
42	88	2.6	2318	2	A45306	notch 3 protein - mou	1.15e+01
43	89	2.6	2524	2	A35844	notch protein - Afric	8.63e+00
44	86	2.5	120	5	17AM	hiv-1 matrix protein	2.03e+01
45	86	2.5	494	2	G64382	acetoacetate synthase	2.03e+01

## ALIGNMENTS

RESULT 1  
 ENTRY glycosyl-phosphatidylinositol-linked protein GDNFR-alpha -  
 TITLE rat  
 ORGANISM  
 DATE 12-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 13-Mar-1998

ACCESSIONS  
 REFERENCE  
 #authors Treanor, J.J.S.; Goodman, L.; de Sauvage, F.; Stone, D.M.; Poulsen, K.T.; Beck, C.D.; Gray, C.; Armanini, M.P.; Pollock, R.A.; Hefti, F.; Phillips, H.S.; Goddard, A.; Moore, M.W.; Bui-Bello, A.; Davies, A.M.; Asai, N.; Takahashi, M.; Vanden, R.; Henderson, C.E.; Rosenthal, A.

#journal Nature (1996) 382:80-83  
 #title Characterisation of a multicomponent receptor for GDNF.  
 #accession S69080  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-24 #label TRE

SUMMARY  
 #length 24 #molecular\_weight 2573 #checksum 2584

Query Match 5.3%; Score 178; DB 2; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 4.46e-14;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MFLATLYFALPLDLMSAEVSGG 24  
 1 MFLATLYFALPLDLMSAEVSGG 24

RESULT 2  
 ENTRY dnant protein - Mycobacterium smegmatis  
 TITLE formal name Mycobacterium smegmatis  
 ORGANISM  
 DATE 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 20-Mar-1998

ACCESSIONS  
 REFERENCE  
 #authors Salazar, L.; Fsihl, H.; de Rosel, E.; Riccardi, G.; Rios, C.; Cole, S.T.; Takiff, H.E.

#journal Mol. Microbiol. (1996) 20:283-293  
 #title Organization of the origins of replication of the chromosomes of Mycobacterium smegmatis, Mycobacterium lepre and Mycobacterium tuberculosis and isolation of a functional origin from M. smegmatis.

#accession S70987



```

##status      preliminary: nucleic acid sequence not shown;
##molecule-type DNA
##residues    1-397 ##label SAL
##cross-references EMBL:X92503
#note         the nucleotide sequence was submitted to the EMBL Data
               Library, October 1995

GENETICS
#gene          dnan
#start-codon   GTC
CLASSIFICATION #superfamily DNA-directed DNA polymerase III beta chain
SUMMARY        length 397 #molecular-weight 41324 #checksum 6687

Query Match      3.2%; Score 108; DB 2; Length 397;
Best Local Similarity 25.0%; Pred. No. 2,23e-02;
Matches 27; Conservative 27; Mismatches 50; Indels 4; Gaps 4.

Db 172 ESVVLAATDRFLAVRELTWTWTAGDVAALVPAKTL-ABAARAGTGDNOVHALTSGA 230
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
361 OTTTTATTARVKKKPLPGPAGSENEIRTHVLPPCANIQAKLSNNGS-THICL-SDS 418
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

231 SVKGDGLGIRSE-GKRSITRLDAEPPKROLLPAEHTAVATIGVAE 277
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
419 DEGKDGLAGASSHTTKSMAAPSCSLSLPYMLTALALLSVSLAE 466

RESULT 3
ENTRY      509668 #type complete
TITLE      hypothetical protein UL103 - human cytomegalovirus (strain
            AD169)
ORGANISM   #format name human cytomegalovirus, human herpesvirus 5
            host Homo sapiens (man)
DATE       07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
            09-Sep-1997
ACCESSIONS 509668
REFERENCE   S09749
#authors    Chee, M.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.;
            Cerny, R.; Horsnell, T.; Hutchinson III, C.A.; Kouzides,
            T.; Martignetti, J.A.; Predde, E.; Satchwell, S.C.;
            Tomlinson, P.; Weston, K.M.; Barrell, B.G.;
            Curr. Top. Microbiol. Immunol. (1990) 154:125-169
            Analysis of the protein-coding content of the sequence of
            human cytomegalovirus strain AD169.
#cross-references MUID:90265039
#accession 509668
#status     nucleic acid sequence not shown; translation not shown
#molecule-type DNA
#residues   1-249 ##label CHE
#cross-references EMBL:X17403; NID:g959591; PID:g1780882
#note       this sequence was submitted to the EMBL Data Library,
            December 1989
SUMMARY      length 249 #molecular-weight 28636 #checksum 1456

Query Match      3.0%; Score 102; DB 2; Length 249;
Best Local Similarity 38.3%; Pred. No. 1.60e-01;
Matches 18; Conservative 7; Mismatches 20; Indels 2; Gaps 2.

Db 106 CTVISCVCNCLTRKCLHDLQYDAVAVRSFRLHHSARLRICS 152
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
178 CTTSMNEVCN-RRKCHKALRQFDPKPAKSYG-KLFCSCRDIACT 222

RESULT 4
ENTRY      570983 #type complete
TITLE      dnan protein - Mycobacterium tuberculosis
ORGANISM   #format name Mycobacterium tuberculosis
DATE       12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change
            20-Mar-1998
ACCESSIONS 570983
REFERENCE   S70980
#authors    Salazar, L.; Fsihl, H.; de Rossi, E.; Riccardi, G.; Rios, C.;
            Cole, S.T.; Takiff, H.E.;
            Mol. Microbiol. (1996) 20:283-293

```

```

#title
Organization of the origins of replication of the chromosome
of Mycobacterium smegmatis, Mycobacterium leprae and
Mycobacterium tuberculosis and isolation of a functional
origin from M. smegmatis.

#accession
S70983

#status
preliminary; nucleic acid sequence not shown;
translation not shown

#molecule_type
DNA

#residues
1-399 ##label SAL

#cross-references
EMBL:X92504
the nucleotide sequence was submitted to the EMBL Data
Library, October 1995

GENETICS

#gene
dnan

CLASSIFICATION
#superfamily
DNA-directed DNA polymerase III beta chain

SUMMARY
#length
399 #molecular_weight
4173 #checksum
9406

Query Match
Best Local Similarity 27.7%: Pred. No. 1,16e-01;
Matches 28; Conservative 27; Mismatches 42; Indels 4; Gaps 4;

Db 178 ATDSRLAVLEWSSAPDIEAALVPARTL-AEAKAGIGSDYRLSIGTGGYKGL 236
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 368 TTARVKNKPLPGSPAGSENIPTHTVLPNCANLQKXKSNVSS-THLCL-SDSDEGKDL 425
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 LGISGN-GKRSITRLDDEPFKROLLPETHVATMDVAE 276
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 426 AGASSHTTRKSMAPPCSLSSPLVMTLALALLSVSLAE 466
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
ENTRY S41819 #type complete
TITLE nucleoporin p62 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
10-Sep-1997
S41819: 543365
REFERENCES S41819
#authors Carmo-Ponseca, M.; Kern, H.; Hurt, E.C.
#journal Eur. J. Cell Biol. (1991) 55:17-30
#title Human nucleoporin p62 and the essential yeast nuclear pore
protein NSP1 show sequence homology and a similar domain
organization.
#accession S41819
##molecule_type mRNA
##residues 1-522 ##label CAR
##cross-references EMBL:X58521
S43365
REFERENCE S43365
#authors Hurt, E.C.
#submission submitted to the EMBL Data Library, March 1991
#accession S43365
##molecule_type mRNA
##residues 1-417, 'RA', 420-430, 'O', 432-506, 'V', 508-522 ##label HUR
##cross-references EMBL:X58521; NID:8432653; RID:8432654
SUMMARY #length 522 #molecular_weight 53268 #checksum 4396

Query Match
Best Local Similarity 31.6%: Pred. No. 3.03e-01;
Matches 18; Conservative 19; Mismatches 19; Indels 1; Gaps 1;

Db 258 LKAPGAAGSGTSTTST-AATATATTTTSSGTGFAINLKLPLAPAGIPSTAAAYAP 313
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 338 LKNAIOAFNGSGSDVTWQAPAPVQTTATTATTAFVKNKRLPGASSENEDIPHTVLP 394
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
ENTRY S60024 #type complete
TITLE bialykynin B1 receptor - rabbit
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
rabbit
DATE 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change
10-Sep-1997
S60024
ACCESSIONS

```

```

ENTRY          JC1121      #type complete
TITLE          leukocyte adhesion protein beta chain precursor - bovine
ALTERNATE_NAMES
ORGANISM       Bos primigenius taurus #common_name cattle
DATE           30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
               05-Sep-1997
ACCESSIONS     JC1121
REFERENCE      Shuster, D.E.; Bosworth, B.T.; Kehrl Jr., M.E.
               #journal 114:267-271
               Gene (1992)
               Sequence of the bovine CD18-encoding cDNA: Comparison with
               the human and murine glycoproteins.
               #cross-references MIMD:92290287
               #accession JC1121
               #molecule_type mRNA
               #residues 1-769 #label SHU
               #cross-references GB:M81233; NID:g162818; PID:g162819
COMMENT         The leukocyte adhesion proteins are noncovalently linked
               heterodimers of distinct alpha and identical beta chains. These
               structurally related glycoproteins mediate cell-adhesion
               reactions, and a deficiency of them is attributed to a genetic
               defect in the expression or structure of their common beta chain.
               #superfamily integrin beta chain.
               #cell_adhesion: cytoskeleton; duplication; glycoprotein;
               heterodimer; leukocyte; pyroglutamic acid; transmembrane
               protein
FEATURES
1-22           #domain signal sequence #status predicted #label SIG\
23-769        #product leukocyte adhesion protein beta chain #status
               predicted #label MAT\
23-670        #domain extracellular #status predicted #label EXT\
445-631       #region cysteine-rich\
701-723       #domain transmembrane #status predicted #label TM\
724-769       #domain intracellular #status predicted #label CYT\
23            #modified site pyrrolidone carboxylic acid (Gln) (1n
               mature form) #status predicted\
50,116,254,501,
642           #binding_site carbohydrate (Asn) (covalent) #status
               predicted
SUMMARY        #length 769 #molecular_weight 84399 #checksum 9562
               #length 2,88; score 96; DB 1; length 769;
               Best Local Similarity 30.88; Pred. No. 1,06e+00;
               Matches 24; Conservative 20; Mismatches 22; Indels 12; Gaps 8;
Db 477 IGKCECOTGGRSSQELGSGSRKDNSS-II-CS-GLADICIGGCVCHTS-----D-VPNK 527
Oy 148 ISKNNCLDAKAC-NLDDTCCKYRRAVITPCTTSMNSNEVCNRKCKALRPFDDKVPAPK 206
Db 528 KIYGQ-FCECDNVNC-ER 543
Oy 207 HSYGMLFCSCRDIACTER 224
RESULT         9
ENTRY          A60385
TITLE          #type complete
ORGANISM       monocyle surface antigen MS2 precursor - mouse
DATE           03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change
               13-Mar-1998
ACCESSIONS     A60385
REFERENCE      Yoshida, S.; Setoguchi, M.; Higuchi, Y.; Akizuki, S.;
               #journal 13:1-10
               Yamamoto, S.
               Int. Immunol. (1990) 2:585-591
               Molecular cloning of cDNA encoding MS2 antigen, a novel cell\
               surface antigen strongly expressed in murine monocytic
               lineage.
               #accession A60385
               #molecule_type mRNA
               #residues 1-826 #label YOS
               #cross-references EMBL:X13335
CLASSIFICATION #superfamily disintegrin homology

```

KEYWORDS	glycoprotein; surface antigen; transmembrane protein
FEATURE	
1-14	#domain signal sequence #status predicted #label SIG\
402-484	#domain disintegrin homology #label DIS\
659-683	#domain transmembrane #status predicted #label TM\
330	#active_site Glu #status predicted
SUMMARY	#length 826 #molecular-weight 89896 #checksum 2686
Query Match	2.8%; Score 96; DB 2; Length 826;
Best Local Similarity 31.4%;	Pred. NO. 1.06e+00;
Matches 16;	Conservative 12; Mismatches 17; Indels 6; Gaps 6;
Db	585 ELYVIGTGC-EERGKVC-MDSCDCLR-VYRSEKSCAKNHNHCNKRCH 632
166 EHSIGNNCIDAAKACNLDDTCRKRYRATITP-CTTS-MSNEVCN-RRCH 193	
RESULT 10	
ENTRY	S33821
TITLE	#type complete
DESCRIPTION	median body protein - Giardia lamblia
TAXID	#formal_name Giardia lamblia
DATE	06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
ACCESSIONS	S33821
REFERENCE	S33821
#authors	Marshall, J.; Holberton, D.V.
#journal	J. Mol. Biol. (1993) 231:521-530
#title	Sequence and structure of a new coiled coil protein from a microtubule bundle in Giardia.
#accession	S33821
#status	Preliminary
#molecule_type	mRNA
#residues	1-857 #label MAR
#cross-references	EMBL:X64517; NID:g312670; PID:g312671
SUMMARY	#length 857 #molecular-weight 100583 #checksum 6805
Query Match	2.8%; Score 94; DB 2; Length 857;
Best Local Similarity 33.3%;	Pred. NO. 1.96e+00;
Matches 15;	Conservative 16; Mismatches 11; Indels 3; Gaps 3;
Db	482 KENSYNFDQLLEOKQMRSDNLAKRAADYE-RVDRRLRLNDKE 525
56 KETNFSILTSGLEARDCEKRSAMELKQKSL-YNCRCKRGMK-REKN 98	
RESULT 11	
ENTRY	S72229
TITLE	#type complete
DESCRIPTION	meiotic recombination protein met-218 - fruit fly (Drosophila melanogaster)
TAXID	#formal_name Drosophila melanogaster
DATE	04-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 26-Feb-1998
ACCESSIONS	S72229
REFERENCE	S72229
#authors	McKim, K.S.; Dahmus, J.B.; Hawley, R.S.
#journal	Genetics (1996) 144:215-228
#title	Cloning of the Drosophila melanogaster meiotic recombination gene met-218: a genetic and molecular analysis of interval 15E.
#accession	S72229
#status	preliminary; nucleic acid sequence not shown
#molecule_type	DNA
#residues	1-1186 #label MCR
#cross-references	EMBL:U35631; NID:g1017733; PID:g1017734
GENETICS	
#gene	met-218
#introns	45/2; 558/3; 1003/2; 1081/1
SUMMARY	#length 1186 #molecular-weight 133551 #checksum 4925
Query Match	2.8%; Score 94; DB 2; Length 1186;
Best Local Similarity 27.0%;	Pred. NO. 1.96e+00;
Matches 20;	Conservative 23; Mismatches 25; Indels 6; Gaps 5;

```

Db      636 FDOOSG-I-INSNSG-IPVEIOINSNSQSVFRASL-PCGCVTMMDFTSPPPIQS 694
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Oy      345 FGNSSDVTM-QRP--PVOTTAITTT-TAFRYKNRP-LGRGSENEIPTHLPCCANLQ 399
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Db      695 SSRUPSSISGPTOL 708
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Oy      400 AQRKLSNVSGSTHL 413
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

RESULT 12
ENTRY   #type fragment
TITLE   hypothetical protein 100-plus (rpoA2 3' region) -
         Thermoplasma acidophilum (fragment)
ORGANISM #formal_name Thermoplasma acidophilum
DATE     12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
         09-Sep-1997

ACCESSIONS
REFERENCE
#authors Klenk, H.P.; Renner, O.; Schwass, V.; Zillig, W.
#journal Nucleic Acids Res. (1992) 20:5226
#title   Nucleotide sequence of the genes encoding the subunits H, B,
         A' and A'' of the DNA-dependent RNA polymerase and the
         initiator tRNA from Thermoplasma acidophilum.

#accession S26728
#status     nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues   1-100 #label KLE
#cross-references EMBL:X68199; NID:g48089; PID:g48097
#experimental_source DSM 1728
#note       the nucleotide sequence was submitted to the EMBL Data
         library, September 1992
         #length 100 #checksum 201

SUMMARY
Query Match      2.7%; Score 92; DB 2; Length 100;
Best Local Similarity 31.5%; Pred. No.3,57e+00;
Matches 17; Conservative 15; Mismatches 20; Indels 2; Gaps 2;

Db      409 GATHTICLNQ-FACTG-ADRSFPASSEALDPPEPLPLPLAKRYNGEVLII 55
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Oy      409 GSTRHCLSDSPFGKDGAGASHTTKRSMAPPSCLSLPLVLMTLTAALLSV 462
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

RESULT 13
ENTRY   #type complete
TITLE   probable membrane protein YNL058c - yeast (Saccharomyces
         cerevisiae)
ALTERNATE_NAMES hypothetical protein N2433; hypothetical protein YNL621
ORGANISM #formal_name Saccharomyces cerevisiae
DATE     16-Feb-1996 #sequence_revision 23-Feb-1996 #text_change
         12-Dec-1997

ACCESSIONS
REFERENCE
#authors Berger, P.; Doljnon, F.; Crouzet, M.
#journal Yeast (1995) 11:967-974
#title   The sequence of a 44 420 bp fragment located on the left arm
         of Chromosome XIV from Saccharomyces cerevisiae.

#accession S58719
#status     nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues   1-336 #label BER
#cross-references EMBL:U01214; NID:g1314216; PID:g994828
#note       the nucleotide sequence was submitted to the EMBL Data
         Library, July 1994

REFERENCE
#authors Berger, P.; Doljnon, F.; Crouzet, M.
#submission submitted to the Protein Sequence Database, April 1996
#accession S62986
#molecule_type DNA
#residues   1-336 #label BEW
#cross-references EMBL:Z71334; NID:g1301928; PID:e239896; PID:g1301929;
         MIPS:YNL058c
#experimental_source strain S288C

GENETICS

```

[illegible]

```

Oy      293  IGTWTFNPNYDSSLSVAPKPCDSCNSGNGDLEDCLK 327
RESULT  15
ENTRY   D70030 #type complete
TITLE   alkalal monooxygenase homolog yvbt - Bacillus subtilis
ORGANISM
DATE    05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
26-Feb-1998
ACCESSIONS
REFERENCE D70030
#authors  AB69580

Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bartero, M.G.; Bessières, P.;
Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Diesterheft, A.;
Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gallier, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Goldlighty, E.J.; Grandi, G.
Giseppli, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.
Kasahara, Y.; Klier-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly
V.; Ogasawa, K.; Ogihara, A.; Oudeg, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetalle, D.; Frowell, S.; Prescott,
A.M.; Pressac, E.; Pujic, P.; Punelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scallion, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekorska, A.; Seror, S.J.; Serron, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Taketani, K.; Takeuchi, M.; Tanakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weltzenger, T.;
Winertz, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.;
Yoshikawa, H.; Danchin, A.
#journal Nature (1997) 390:249-256
#title The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#accession D70030
#status preliminary; nucleic acid sequence not shown:
#molecule_type DNA
#residues 1-336 #label KUN
#experimental_source strain 168
GENETICS
#gene yvbt
CLASSIFICATION
#superfamily yvbt protein
SUMMARY #length 336 #molecular_weight 37101 #checksum 2405

Query Match 2.7%; Score 91; DB 2; Length 336;
Best Local Similarity 40.0%; Pred. No. 4.80e+00;
Matches 16; Conservative 12; Mismatches 10; Indels 2; Gaps 2;

Db 132 NSGDEPFOLEELNRYKPSGYNRNOVRAIPGSDVPM 171
||||: | | | | | : : : | | | : |
OY 317 NSGNDLEDCKLFLNFKDKWTCLKNAIQAFNGSGSVTW 354

Search completed: Tue Jun 23 18:34:22 1998
Job time : 121 secs.

```

THIS PAGE BLANK (USPTO)

\*\*\*\*\*  
 WISE (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 on: Tue Jun 23 18:28:42 1998; Mspar time 16.07 Seconds  
 730.390 Million cell updates/sec  
 Molecular output not generated.

Title: >US-08-866-354-4  
 Description: (1-468) from US08866354.pep  
 Perfect Score: 3384  
 Sequence: 1 MFLATLYFALPLDLMSAE.....PVIMLTRALALVLSAETS 468

Scoring table: PAM 150  
 Gap 11

Searched: 69111 segs, 25083644 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot35  
 1:swiss1

Statistics: Mean 47.975; Variance 67.645; scale 0.709

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

No.	Score	Query	Match	Length	ID	Description	Pred. No.
1	3384	100.0	468	1	GDNR_RAT	GDNR RECEPTOR ALPHA PR	0.00e+00
2	3335	98.6	468	1	GDNR_MOUSE	GDNR RECEPTOR ALPHA PR	0.00e+00
3	3187	94.2	464	1	GDNR_HUMAN	GDNR RECEPTOR ALPHA PR	0.00e+00
4	2598	76.8	469	1	GDNR_CHICK	GDNR RECEPTOR ALPHA PR	0.00e+00
5	1556	46.0	464	1	NRRF_HUMAN	NEUTRURIN RECEPTOR ALP	0.00e+00
6	1539	45.5	463	1	NRRF_MOUSE	NEUTRURIN RECEPTOR ALP	0.00e+00
7	1511	44.7	463	1	NRRF_CHICK	NEUTRURIN RECEPTOR ALP	0.00e+00
8	108	3.2	397	1	DP3B_MYCM	DNA POLYMERASE III, BE	1.15e-03
9	102	3.0	249	1	ULAJ_HCMVA	NUCLEAR PORE GLYCOPROT	2.80e-02
10	100	3.0	522	1	NU62_HUMAN	B1 BRADYKININ RECEPTOR	4.11e-02
11	99	2.9	352	1	BRB1_RABIT	PROTEIN-EXPORT PROTEIN	1.86e-01
12	95	2.8	169	1	SECB_HAEIN	GAG POLYPROTEIN (CONTA	1.28e-01
13	96	2.8	504	1	GAG_HV1MA	CELL SURFACE ADHESION	1.28e-01
14	96	2.8	769	1	ITB2_BOVIN	CELL SURFACE ADHESION	1.28e-01
15	96	2.8	826	1	MS2_MOUSE	CELL SURFACE ADHESION	1.28e-01
16	94	2.8	857	1	MEB_GIALA	CELL SURFACE ADHESION	1.28e-01
17	92	2.7	100	1	YRP4_THEAC	CELL SURFACE ADHESION	1.28e-01
18	91	2.7	316	1	YNEB_YEAST	CELL SURFACE ADHESION	1.28e-01
19	92	2.7	320	1	KITH_PVANS	CELL SURFACE ADHESION	1.28e-01
20	90	2.7	327	1	UPAR_HVIM2	CELL SURFACE ADHESION	1.28e-01
21	91	2.7	388	1	GAG_HVIM2	CELL SURFACE ADHESION	1.28e-01
22	91	2.7	469	1	BFI_HUMAN	CELL SURFACE ADHESION	1.28e-01
23	90	2.7	475	1	SIM1_YEAST	CELL SURFACE ADHESION	1.28e-01

24	91	2.7	503	1	GAG_HV1JR	GAG POLYPROTEIN (CONTA	7.97e-01
25	93	2.7	567	1	CH3_CANAL	CHITINASE 3 PRECURSOR	3.88e-01
26	93	2.7	600	1	SP96_DICDI	SPORE COAT PROTEIN SP9	3.88e-01
27	92	2.7	688	1	CIS_HUMAN	COMPLEMENT C1S COMPONE	5.57e-01
28	93	2.7	695	1	CASP_MESAV	CASP-DEPENDENT SERI	3.88e-01
29	93	2.7	798	1	HEPA_HCMVA	DNA HELICASE/PRINASE C	3.88e-01
30	89	2.6	247	1	SUMT_PSEFL	UROPOPHRYLIN-III C-MET	1.62e+00
31	88	2.6	378	1	LEUK_RAT	LEUCOSTALIN PRECURSOR	2.29e+00
32	88	2.6	415	1	CBPB_RAT	CARBOXYPEPTIDASE B PRE	2.29e+00
33	89	2.6	491	1	CC1_YEAST	CELL DIVISION CONTROL	1.62e+00
34	88	2.6	506	1	GAG_HV1NM	GAG POLYPROTEIN (CONTA	2.29e+00
35	87	2.6	514	1	GUXC_FUSOX	POTATIVE EXOGUCANASE	3.22e+00
36	88	2.6	520	1	GUXC_HV2SB	GAG POLYPROTEIN (CORE	2.29e+00
37	87	2.6	525	1	NU62_RAT	NUCLEAR PORE GLYCOPROT	1.62e+00
38	89	2.6	563	1	ENV_BAEVM	ENV POLYPROTEIN PRECUR	2.29e+00
39	88	2.6	743	1	STB5_YEAST	STB5 PROTEIN	2.29e+00
40	88	2.6	769	1	ITB2_PIG	CELL SURFACE ADHESION	2.29e+00
41	89	2.6	770	1	ITB2_MOUSE	CELL SURFACE ADHESION	2.29e+00
42	88	2.6	1188	1	PSA_METJA	PROBABLE PHOSPHOLIPY	2.29e+00
43	88	2.6	1409	1	HAP1_HAEIN	ADHESION AND PENETRATI	2.29e+00
44	88	2.6	2318	1	NRC3_MOUSE	NEUROGENIC LOCUS NOTCH	1.62e+00
45	89	2.6	2524	1	NOTC_XENLA	NEUROGENIC LOCUS NOTCH	1.62e+00

## ALIGNMENTS

RESULT 1 STANDARD: PRT: 468 AA.

ID	GDNR_RAT	STANDARD:	PRT:	468 AA.
AC	062997:			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	GDNR RECEPTOR ALPHA PRECURSOR (GDNR-ALPHA) (TGF-BETA RELATED			
DE	NEUROTROPHIC FACTOR RECEPTOR 1).			
GN	GDNFRA OR TRNRI.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-RETINA;			
RX	MEDLINE: 96270513.			
RA	JING S., MEN D., YU Y., HOLST P.L., LUD Y., FANG M., TAMIR R.,			
RA	ANTONIO L., HU Z., CUPPLES R., LOUIS J.-C., HU S., ALTHROCK B.W.,			
RA	FOX G.M.,			
CC	CELL 85:1113-1124(1996).			
CC	- FUNCTION: RECEPTOR FOR GDNF. MEDIATES THE GDNF-INDUCED			
CC	AUTOPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR.			
CC	- SUBUNIT: 2 MOLECULES OF GDNF-ALPHA ARE THOUGHT TO FORM A COMPLEX			
CC	WITH THE DISULFIDE-LINKED GDNF DIMER AND WITH 2 MOLECULES OF RET.			
CC	- TISSUE SPECIFICITY: EXPRESSED IN LIVER, BRAIN AND KIDNEY.			
CC	- SIMILARITY: BELONGS TO THE GDNF FAMILY.			
DR	EMBL: U59486; G1399863; -			
KW	RECEPTOR; GLYCOPROTEIN; GPI-ANCHOR; MEMBRANE; SIGNAL.			
FT	SIGNAL	1	24	POTENTIAL.
FT	CHNIN	25	?	GDNR RECEPTOR ALPHA.
FT	PROPEP	?	468	HYDROPHOBIC, REMOVED DURING MATURATION
FT	DOMAIN	362	369	POLY-THR.
FT	CARBOHYD	59	59	POTENTIAL.
FT	CARBOHYD	347	347	POTENTIAL.
FT	CARBOHYD	406	406	POTENTIAL.
SO	SEQUENCE	468 AA;	6A7AAB2A	CRC32;

Query Match 100.0%; Score 3384; DB 1; Length 468;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

D	b	61	SLTGLEKDCBCSRAMELKO	SLYNCRCKRGKKKENCN	RIYMSWQ	SLQGDLLDSD	120		
Q	y	61	SLTSGLEKDCBCSRAMELKO	SLYNCRCKRGKKKENCN	RIYMSWQ	SLQGDLLDSD	120		
D	b	121	YEPNRSRLSDIEFRAVP	ISDFVQOOVEHIS	KGNNCLDAKACN	LDPTCKKYSRAYITPCT	180		
Q	y	121	YEPNRSRLSDIEFRAVP	ISDFVQOOVEHIS	KGNNCLDAKACN	LDPTCKKYSRAYITPCT	180		
D	b	181	SMSNEVCNRKCKHALQ	FOFDKVPAKHSYGM	LFECSCD	IACTERRROTIVPVC	240		
Q	y	181	SMSNEVCNRKCKHALQ	FOFDKVPAKHSYGM	LFECSCD	IACTERRROTIVPVC	240		
D	b	241	PNCISLSDSC	TNYICGRSLAD	FTFNQPSRSVSN	CKENYADCLLAYSG	LIGYWT	300	
Q	y	241	PNCISLSDSC	TNYICGRSLAD	FTFNQPSRSVSN	CKENYADCLLAYSG	LIGYWT	300	
D	b	301	YVDSSSLVAPMC	CCSNGNDLEBCL	KFLNF	KONCTLNAIOAF	NGSVYTMOPAPV	360	
Q	y	301	YVDSSSLVAPMC	CCSNGNDLEBCL	KFLNF	KONCTLNAIOAF	NGSVYTMOPAPV	360	
D	b	361	QTTATATTTAA	RVRNKP	LGPA	GSENEIPTHYL	PPCANIQAKL	SNVSGSTHLC	420
Q	y	361	QTTATATTTAA	RVRNKP	LGPA	GSENEIPTHYL	PPCANIQAKL	SNVSGSTHLC	420
D	b	421	GKGDLAGASS	HIITTKSMAAP	SCSLSLP	VLMTAL	TALAL	SVSLAETS	468
Q	y	421	GKGDLAGASS	HIITTKSMAAP	SCSLSLP	VLMTAL	TALAL	SVSLAETS	468

RESULT	2	STANDARD	PRT	468 AA.
ID	GNDR_MOUSE			
AC	P97785:			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	GNDF RECEPTOR ALPHA PRECURSOR (GNDF-ALPHA) (TGF-BETA RELATED NEUROTROPHIC FACTOR RECEPTOR 1).			
GN	GNDFRA OR TRNRL.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; EUTHERIA; RODENTIA.			
RP	[1]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE-DORSAL ROOT GANGLION;			
RA	MATLAB K.;			
RL	SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDJ DATA BANKS.			
CC	-1- FUNCTION: RECEPTOR FOR GDNF. MEDIATES THE GDNF-INDUCED AUTOPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR (BY SIMILARITY).			
CC	-1- SUBUNIT: 2 MOLECULES OF GDNF-ALPHA ARE THOUGHT TO FORM A COMPLEX WITH THE DISULFIDE-LINKED GDNF DIMER AND WITH 2 MOLECULES OF RET (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE GDNF FAMILY.			
DR	EMBL: AB000800; GI816442; -			
KW	RECEPTOR; GLYCOPROTEIN; GPI-ANCHOR; MEMBRANE; SIGNAL.			
FT	SIGNAL	1	24	POTENTIAL.
FT	CHAIN	25	7	GNDF RECEPTOR ALPHA.
FT	PROPEP	?	468	HYDROPHOBIC, REMOVED DURING MATURATION (POTENTIAL).
FT	DOMAIN	362	369	POLY-THR.
FT	CARBOHYD	59	59	POTENTIAL.
FT	CARBOHYD	347	347	POTENTIAL.
FT	CARBOHYD	406	406	POTENTIAL.
SO	SEQUENCE	468 AA;	51782 MW;	6C64C182 CRC32;
Query Match		98.6%;	Score 3335;	DB 1; length 468;
Best Local Similarity		97.4%;	Pred. No. 0.00e+00;	
Matches	466;	Conservative	10;	Mismatches 2; Indels 0; Gaps 0.

Qy	1	MFATLYALFALDLDLMSAEVSGGDRDLCVKASQJCKEDSGSCSKYTTLQCVAGKRTNF	60
Db	61	SLTSGLEAKDEDCRAMEALNOKSLYNCRCRKGMKERNCLRIYWSMYOSLOGNDLLEDSP	120
Qy	61	SLTSGLEAKDEDCRAMEALNOKSLYNCRCRKGMKERNCLRIYWSMYOSLOGNDLLEDSP	120
Db	121	YEPVNSRLSDIFRAVPFISDVFOQVEHISKGNNCLDAKACNDDTCKKYSAYITPCTT	180
Qy	121	YEPVNSRLSDIFRAVPFISDVFOQVEHISKGNNCLDAKACNDDTCKKYSAYITPCTT	180
Db	181	SMSNEVCNRKRCCHKALROFEDKVPAKHSYGMLFCSCHDVACTERRROTIVPVCSYEERER	240
Qy	181	SMSNEVCNRKRCCHKALROFEDKVPAKHSYGMLFCSCHDVACTERRROTIVPVCSYEERER	240
Db	241	PNCNLADSCCTNICYRSRLADFTTNCPBSRSYSNCLKENYADCLLAYIGLGITVWTPN	300
Qy	241	PNCNLADSCCTNICYRSRLADFTTNCPBSRSYSNCLKENYADCLLAYIGLGITVWTPN	300
Db	301	YIDSSLSVAAWCCCSNSGNDLEBCLFLNFKRNTCLKNAIQAFGSGSVTMMQAPPV	360
Qy	301	YIDSSLSVAAWCCCSNSGNDLEBCLFLNFKRNTCLKNAIQAFGSGSVTMMQAPPV	360
Db	361	OTTAMTTTARINKNPLGPAGSSENEIPTHLPPCANLQAOXKLSNVSSTHILCLSDNDY	420
Qy	361	OTTAMTTTARINKNPLGPAGSSENEIPTHLPPCANLQAOXKLSNVSSTHILCLSDNDY	420
Db	421	GRDGLAGASSHITTKSMAAPSCGLSSLPVAVFALALALSVSLAETS	468
Qy	421	GRDGLAGASSHITTKSMAAPSCGLSSLPVAVFALALALSVSLAETS	468

RESULT	3	STANDARD:	PRT:	464 AA.
AC	PS6159:			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	GDNF RECEPTOR ALPHA PRECURSOR (GDNFR-ALPHA) (TGF-BETA RELATED			
DE	NEUROTROPIC FACTOR RECEPTOR 1).			
GN	GDNFR OR TRNRI.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-SUBSTANTIA NIGRA;			
RX	MEDLINE: 96270513.			
RA	JING S., WEN D., YU Y., HOLST P.L., LUO Y., FANG M., TAMIR R.,			
RA	ANTONIO L., HU Z., CUPPLES R., LOUIS J.-C., HU S., ALTRICK B.W.,			
RA	FOX G.M.;			
RL	CELL 85:1113-1124(1996).			
CC	-1- FUNCTION: RECEPTOR FOR GDNF. MEDIATES THE GDNF-INDUCED			
CC	ANTIPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR (BY			
CC	SIMILARITY).			
CC	-1- SUBUNIT: 2 MOLECULES OF GDNFR-ALPHA ARE THOUGHT TO FORM A COMPLEX			
CC	WITH THE DISULFIDE-LINKED GDNF DIMER AND 2 MOLECULES OF RET (BY			
CC	SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (BY			
CC	SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE GDNFR FAMILY.			
CC	MIM: 601496; -.			
DR	RECEPTOR; GLYCOPROTEIN; GPI-ANCHOR; MEMBRANE; SIGNAL.			
KW	SIGNAL	1	24	POTENTIAL.
FT	CHAIN	25	?	GDNF RECEPTOR ALPHA.
FT	PROPEP	?	464	HYDROPHOBIC, REMOVED DURING MATURATION
FT				(POTENTIAL).
FT	POLY-THR.			POTENTIAL.
FT	CARBOHYD	361	368	POTENTIAL.
FT	CARBOHYD	59	58	POTENTIAL.
FT	CARBOHYD	346	346	POTENTIAL.
FT	CARBOHYD	405	405	POTENTIAL.
SO	SEQUENCE	464 AA;	51291 MF;	2C8C3574 CRC32;
Query Match		94.24;	Score 3187;	DB 1; Length 464;





FT	CARBOHYD	52	52	POTENTIAL.
FT	CARBOHYD	357	357	POTENTIAL.
FT	CARBOHYD	413	413	POTENTIAL.
SO	SEQUENCE	463 AA;	51598 NM;	0A2165C0 CRC32;
Query Match 45.5%; Score 1539; DB 1; Length 463;				
Best Local Similarity 49.1%; Pred. No. 0.00e+00;				
Matches 214; Conservative 96; Mismatches 113; Indels 13; Gaps 1				
Db	37 QVDCRANLCAEAESNCSSRYRTLRQCLAGDRN--TW--LANK-ECQALAEVLOESPLY	91		
Oy	26 RLDCVRKASDQCLKEGSCSTFYRTLRQCVAGKETNSLTSLEAKDECRSAEALQKSL	85		
Db	92 DCRKRKMKELQCLQIYNSIHGLTEGEFEASPYEYYSRLSDIRLASISFGICAD	151		
Oy	86 NCRKRMRKMKELKCLTIYMSYQSL-QGNDLLEDSPEYEVNSRLSDIRAVPEFTSDVFOQ	144		
Db	152 PYVSAAGNCHCLDAKAKCNLMDNCKKTRSSYISICNREISPECRNRRCHKALRQFEDRY	211		
Oy	145 VEHSKRNCLDAKAKCNLDIDCKTKRSYITPCTTSHS-NEVCRRKCHALRQFEDRY	203		
Db	212 PSEYTYRMLFCSCQDOACAEERRRQITLPSGSYDEKEKPCNDLRSLCTHDLCSRLADF	271		
Oy	204 PAKHSYEMLCSCORDIACETERRRQITLPSGSYDEKEKPCNDLRSLCTHDLCSRLADF	263		
Db	272 HANCRASYRTITSCPADNTQACLSAGAGIGEDMTPNYVDSIPGICIVYSPMCNCRGSGNM	331		
Oy	264 FTNCPSPRSYSNCLKENYADCLAYSGLIGITVMPNYVDS-S-LSYAPCDSNSGND	321		
Db	332 EEECEKLFKFTENPCLRNIAFGAGTGVN-SPKGFETSAQPAVE-KTPSLPDLDS	389		
Oy	322 LEDCLAKLNFKNCTCLKNHIAQFNGSVYTMQAPRQITATTTAFKYNKKPLGPA	381		
Db	390 DSTS-LGTSYITCTSIQEOGLKANSKELSMCFELTITNISPGSKYIKLYSGSCARAT	448		
Oy	382 GSENEIFTHVLPFCANLQAGKLSNVSGSTHCLSDSPDGKGLGAGSHITTKSMAP	441		
Db	449 STALTALPLLMVT-LA	463		
Oy	442 SCSLSLPLVLMITALA	457		
RESULT 7				
ID	NRTR-CHICK	STANDARD:	PRT:	465 AA.
AC	013157;			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	NEURURIN RECEPTOR ALPHA PRECURSOR (NTNR-ALPHA) (NTRNR-ALPHA) (GDNF RECEPTOR BETA) (GDNFR-BETA).			
GN	GDNFR.			
OC	GALLUS GALLUS (CHICKEN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;			
OC	GALLIFORMES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	BUJ-BELLO A., ADU J., PINON L.G., HORTON A., THOMPSON J.,			
RA	ROSENTHAL A., CHINGCHETRU M., BUCHANAN V.L., DAVIES A.M.;			
RL	NATURE 387:721-724(1997).			
CC	-1- FUNCTION: RECEPTOR FOR NEURTURIN. MEDIATES THE NRTN-INDUCED			
CC	AUTOPHOSPHORYLATION AND ACTIVATION OF THE RET RECEPTOR.			
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (BY			
CC	SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE GDNFR FAMILY.			
DR	EMBL: U09542; G2213805; -.			
CC	RECEPTOR: GLYCOPROTEIN; GPI-ANCHOR: MEMBRANE; SIGNAL.			
KM	RECEPTOR: GLYCOPROTEIN; GPI-ANCHOR: MEMBRANE; SIGNAL.			
FT	SIGNAL	1	?	POTENTIAL.
FT	CHAIN	?	?	NEURURIN RECEPTOR ALPHA.
FT	PROPER	?	465	HYDROPHOBIC, REMOVED DURING MATURATION
FT	CARBOHYD	355	355	POTENTIAL.
FT	CARBOHYD	387	387	POTENTIAL.
FT	CARBOHYD	412	412	POTENTIAL.

SQ	SEQUENCE	465 AA;	51908 MW;	SCA073E4 CRC32:	
	Query Match	44.7%;	Score 1511;	DB 1; Length 465;	
	Best Local Similarity	46.6%;	Pred. NO. 0.00e+00;		
	Matches	203;	Conservative 102;	Mismatches 116; Indels 15; Gaps 11.	
Db	38	VDCIRANKLCAAEAGSSRRYRTLRQCCLAGRBRN--TM---LNKK-ECQAALVLDSEPLYD	92		
Oy	27	LDCVAAASOQCKEBOSCSTKRYRPLRCVCAGKETNSLVISGLAEKACRCRAMBALOKSLYN	86		
Db	93	CRCRKGRKEIOCLQVYYSIHILGAEGEFEYEAPEYEPITRLSDIFRLASFSGM-DPA	151		
Oy	87	CRCRKGMKREKNCLRIVYSMYQS-LQGNDLEDSEYEVEVNRSLDIFRAVFIDVQY	145		
Db	152	TN-SKSNHCLDAAKACNLNDCKRLRSYISTCSKEISATEHCSCRKCHALROFPDNP	210		
Oy	146	EHSIGNNCLDLAAKACNLDDCTKKRSAYIPTCTSMSEN-EVCNKRKCCHKALROFPDPVP	204		
Db	211	SEYTRLLFGCCKDOACAEPROTIVPCEVDENKPKMCLDRANCRADHLCSRBLDFH	270		
Db	205	AKHSTGMLFCCSRDIACETERRQOTIVPACSIEBERPNCLSLDQCKNTYCRLSRLDFE	264		
Oy	271	ANQAFOSFLTSCPEDNYOACIGSYTGLIGDMTPNYADASTTSITSPMSCSCAGSLE	330		
Oy	265	TNGCESHSVSNCLEKENYADCLLAWSGLIGTMPENVYDS--SLSVAPWCDCNSGNDL	322		
Db	331	EECEFFLDEFTENPCRLRAIQAFNGCTGVNL-SPKNSEPPIITMLPKVE-KSPALPDIND	388		
Oy	323	EDCLFLNFEPDNCLKAIAFGSGSVTMQAPRPQTATATTATTAFRVKNRPLGPAG	382		
Db	369	SNMTYDTSIITTCSTIOEHGOKLNKSQSLCYESTOLTPTDMPQCKFYVOKAGSHR	448		
Oy	363	SENEPTHYLPFCANLQAKLKSNVSGSTHLCLSPSDKRGKLACASSHTTKSMAPPS	442		
Db	449	AARILPAVPVYLKLL	464		
Oy	443	CS--LSSLPLVLMLTL	456		
RESULT	ID	DP3B_MYCSM	STANDARD;	PRT;	397 AA.
AC	P52851;				
DT	01-OCT-1996 (REL. 34, CREATED)				
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	DNA POLYMERASE III, BETA CHAIN (BC 2.7.7.7).				
GN	DNAM.				
GC	MYCOBACTERIUM SMEGNATIS.				
	PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.				
	[1]				
RC	SEQUENCE FROM N.A.				
RX	STRAIN=MC26;				
RA	MEDLINE: 96310367.				
RA	SALAZAR L., FSIHI H., DE ROSSI E., RICCARDI G., RIOS C.,				
RA	COLE S.T., TAKIFE H.E.;				
RL	MOL. MICROBIOL. 20:283-293(1996).				
CC	-1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME				
CC	RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.				
CC	THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.				
CC	THE BETA CHAIN IS REQUIRED FOR INITIATION OF REPLICATION ONCE IT				
CC	IS CLAMPED ONTO DNA. IT SLIDES FREELY (BIDIRECTIONAL AND ATP-				
CC	INDEPENDENT) ALONG DUPLICATION DNA (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -				
CC	N PYROPHOSPHATE + DNA(N).				
CC	-1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND				
CC	THETA CHAINS) THAT CAN REPAIR SHORT GAPS CREATED BY NUCLEASE				
CC	IN DUPLEX DNA. FOR EFFICIENT REPLICATION OF THE LONG, SINGLE-				
CC	STRANDED TEMPLATES, POL III REQUIRES THE AUXILIARY CHAINS BETA,				
CC	GAMMA, AND DELTA.				
EMBL:	X92503: E208988: -				
XR	DNA-DIRECTED DNA POLYMERASE: DNA REPLICATION.				
SW	SEQUENCE 397 AA: 41324 MW: EEFB3064 CRC32:				

```

Query Match      3.2% Score 108: DB 1: Length 397:
Best Local Similarity 25.0% Pred. No. 1.15e-03:
Matches 27: Conservative 27: Mismatches 50: Indels 4: Gaps 4:

Db 172 ESVVLAAIDRRLVRLVETLWTWTAGDVAEVLVPAKTL-AEAAKAGTDGNQVHLAAGSA 230
Oy 361 OTTATTTTARVANKKPLGPAAGSENELEPTHLPPCALQKQKLKSNVSGS-THLCL-SDS 418
Db 231 SVGKDGDLGIRSE-GKRSTTRLLDAEFPKFRQLLPAAETAVATIGVAE 277
Oy 419 DEGKDGLAGASSHTTKSMAPPSCSLSPLVLTLLAALLSVSLAE 466

RESULT 9
ID ULA3_HCMVA STANDARD; PRT; 249 AA.
AC P16734;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE PROTEIN UL103.
GN UL103.
OS HUMAN CYTOMEGALOVIRUS (STRAIN AD169).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.
RN
RP SEQUENCE FROM N.A.
RX MEDLINE; 90269039.
RA CHEE M.S., BANKER A.T., BECK S., BOHNI R., BROWN C.M., CERNY R.,
RA HOSNELL T., HUTCHISON C.A. III, KOZARIDES T., MARGINETTI J.A.,
RA REDDIE E., SATCHEL S.C., TOMLINSON P., WESTON K.M., BARRELL B.G.;
RL CURR. TOP. MICROBIOL. IMMUNOL. 154:125-169(1990).
-1 SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL7,
-1 EHV-1 55, VZV 53, EBV BBRF2, HCMV UL103 AND HSV-1 42.
CC
CC EMBL; X17403; G59711; -
DR PIR; S09868; S09868.
DR
SQ SEQUENCE 249 AA; 28636 MW; 239ADB2D CRC32:

Query Match      3.0% Score 102: DB 1: Length 249:
Best Local Similarity 38.3% Pred. No. 1.28e-02:
Matches 18: Conservative 7: Mismatches 20: Indels 2: Gaps 2:

Db 106 CTYVISCVENCLTRKCLHDLLOYLDVAVNRSFQRLHLHSAARLICS 152
Oy 178 CTTSMNSEVCN-RRCKRALRPFQDKPAKHSYG-MLESCRDIACT 222

RESULT 10
ID NU62_HUMAN STANDARD; PRT; 522 AA.
AC P37198;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE NUCLEAR PORE GLYCOPROTEIN P62 (NUCLEOPORIN P62).
GN NU62.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92007939.
RA CARMO-PONSECA M., KERN H., HURT E.C.;
RA EUR. J. CELL BIOL. 55:17-30(1991).
-1 FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEAR PORE COMPLEX.
-1 THE N-TERMINAL IS PROBABLY INVOLVED IN NUCLEOTYPLASMIC
TRANSPORT. THE C-TERMINAL IS PROBABLY INVOLVED IN PROTEIN-PROTEIN
INTERACTION VIA COILED-COIL FORMATION AND MAY FUNCTION IN
ANCHORAGE OF P62 TO THE PORE COMPLEX.
-1 SUBCELLULAR LOCATION: CENTRAL REGION OF THE NUCLEAR PORE, WITHIN
THE TRANSPORTER. DURING MITOTIC CELL DIVISION, IT ASSOCIATES WITH
THE POLES OF THE MITOTIC SPINDLE.
-1 DOMAIN: CONTAINS X-F-X-F-G REPEATS.
-1 PHY: O-GLYCOSYLATED. CONTAINS ABOUT 10 N-ACETYLGLUSAMINE SIDE
CHAIN SITES PREDICTED FOR THE ENTIRE PROTEIN, AMONGST WHICH ONLY
ONE IN THE C-TERMINAL.

```



CC P04594; (REL. 05, CREATED)  
DT 13-AUG-1987 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DE GAG POLYPROTEIN (CONTAINS: CORE PROTEINS P17, P24, P2, P7, P1, P6).  
GN GAG.  
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (MAL ISOLATE) (HIV-1).  
OC VIRIDAE: SS-RNA ENVELOPED VIRUSES: POSITIVE-STRAND; RETROVIRIDAE;  
OC LENTIVIRINAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 86245056.  
RA ALIZON M., MAIN-HOBSON S., MONTAGNIER L., SONIGO P.;  
RL CELL 46:63-74(1986).  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
DR -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
DR EMBL: K03456; G328021; -.  
DR EMBL: X04415; G60229; -.  
DR EMBL: A07116; G492872; -.  
DR HSPB; P03348; 14YN.  
DR HIV; K03456; GAGSMAL.  
DR AIDS; CORE PROTEIN; POLYPROTEIN; MYRISTYLATION; PHOSPHORYLATION.  
FT INIT\_MET 0  
FT LIPID 1 1  
FT CHAIN 1 137 CORE PROTEIN P17 (MATRIX PROTEIN).  
FT CHAIN 138 368 CORE PROTEIN P24 (CORE ANTIGEN).  
FT CHAIN 369 383 CORE PROTEIN P2.  
FT CHAIN 384 437 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
FT CHAIN 438 453 CORE PROTEIN P1.  
FT CHAIN 454 504 CORE PROTEIN P6.  
SQ SEQUENCE 504 AA: 56001 MW: 4500CD4 CRC32;  
Query Match 2.8%; Score 96; DB 1; Length 504;  
Best Local Similarity 41.9%; Pred. No. 1,286-01;  
Matches 18; Conservative 11; Mismatches 9; Indels 5; Gaps 5  
DB 27 KYR-LKHLVMSRELEFRALNGLLETEGCGQINQSL-QSL 67  
OY 45 KYRTRROCV-AGKETN-FSLTSLG-L-EARDECRSAMEALKRQSL 84  
RESULT 14  
ITB2\_BOVIN STANDARD: PRT; 769 AA.  
CC P32592;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DE CELL SURFACE ADHESION GLYCOPROTEINS LEA-1, CR3 AND P150, 95, BETA-  
DE SUBUNIT PRECURSOR (INTEGRIN BETA-2) (CD18 ANTIGEN).  
GN ITB2 OR CD18.  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA; MAMMALIA;  
OC EUTHERIA; ARTIODACTYLA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 92290287.  
RA SHUSTER D.E., BOSMORTH B.T., KEHRLI M.E. JR.;  
RL GENE 114:267-271(1992).  
RN [2]  
RP VARIANT LAD GUY-128.  
RX MEDLINE: 93028437.  
RA SHUSTER D.E., KEHRLI M.E. JR., ACKERMANN M.R., GILBERT R.O.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 89:9225-9229(1992).  
CC -1- FUNCTION: ASSOCIATES WITH ALPHA-L (LEA-1) TO INTERACT WITH ICBM-1,  
CC AND WITH ALPHA-M (MAC-1) OR ALPHA-X TO FORM THE RECEPTOR FOR THE  
CC ICB FRAGMENT OF THE THIRD COMPLEMENT COMPONENT.  
CC -1- SUBUNIT: DIMER OF AN ALPHA AND BETA SUBUNIT. BETA-2 ASSOCIATES

CC	WITH DIFFERENT ALPHA CHAINS.
CC	-1- SUBCELLULAR LOCATION: TYPE 1. MEMBRANE PROTEIN.
CC	-1- PPM: THE CISTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE BONDS.
CC	-1- DISEASE: DEFECTS IN CD18 ARE THE CAUSE OF LEUCOCYTE ADHESION DEFICIENCY (LAD). THE MUTATION CAUSING LAD (GLY-128) IS PREVALENT AMONG HOLSTEIN CATTLE THROUGHOUT THE WORLD. PLACING THIS DISORDER AMONG THE MOST COMMON GENETIC DISEASES KNOWN IN ANIMAL AGRICULTURE. ALL CATTLE WITH THE MUTANT ALLELE ARE RELATED TO ONE BULL, WHO THROUGH THE USE OF ARTIFICIAL INSEMINATION Sired MANY CALVES IN THE 1950S AND 1960S.
CC	-1- SIMILARITY: WITH OTHER BETA CHAINS FROM THE INTEGRIN FAMILY OF CELL-SURFACE RECEPTOR.
CC	EMBL; M81233; G162819; -.
DR	PIR; JC1121; JC1121.
DR	PROSITE; PS00243; INTEGRIN.BETA. 3.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR	PROSITE; PS01186; EGF_2; UNKNOWN_2.
KW	INTEGRIN; CELL ADHESION; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; EXTRACELLULAR MATRIX; CYTOSKELETON; SIGNAL; DISEASE MUTATION.
KW	EXTRACELLULAR MATRIX; CYTOSKELETON; SIGNAL; DISEASE MUTATION.
FT	SIGNAL 1 22
FT	CHAIN 23 769
FT	DOMAIN 23 700
FT	TRANSMEM 701 723
FT	DOMAIN 724 769
FT	DOMAIN 449 617
FT	REPEAT 449 496
FT	REPEAT 497 540
FT	REPEAT 541 581
FT	REPEAT 582 617
FT	SIZE 397 399
FT	CARBOHYD 50 50
FT	CARBOHYD 116 116
FT	CARBOHYD 254 254
FT	CARBOHYD 501 501
FT	CARBOHYD 642 642
FT	VARIANT 128 128
SO	SEQUENCE 769 AA; 84400 MW; 46F81EE CRC32;
Db	477 IGKNCCEOTGGRSQOELGEGSCRKDNSS-II-CS-GLGDCICGOCVCCHTS-----D-VPNK 527
Oy	148 ISKNNNCILDAKAC-NLDDPCKKRYRSAYITPCTTSMNSNEVCNKRCKHKLROFPDKVPAB 206
Db	528 KIYGO-FCGCDNVCN-ER 543
Oy	207 HSYGMFLFCSCRDIACTER 224
RESULT 15	
AC	MS2 MOUSE STANDARD; PRT; 826 AA.
AC	005910;
DT	01-NOV-1995 (REL. 32, CREATED)
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE	CELL SURFACE ANTIGEN MS2 PRECURSOR (EC 3.4.24.-) (MACROPHAGE CISTEINE RICH GLYCOPROTEIN).
DE	ADAM8 OR MS2.
OS	MUS MUSCULUS (MOUSE).
OC	EUFAROT; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; RODENTIA.
CC	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-ICR;
RA	YAMAMOTO S., YOSHIYAMA K., SETOGUCHI M., MATSUURA K., HIGUCHI Y., AKIZUKI S.;
RA	SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RP	[2]
RP	PRELIMINARY SEQUENCE FROM N.A.
RC	STRAIN-ICR;



\*\*\*\*\*  
[W] [O] [R] [E] [I] (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 23 18:29:51 1998; Mspar time 28.43 Seconds  
Molecular output not generated. 693.273 Million cell updates/sec

Title: >US-08-866-354-4  
Description: (1-468) from US08866354.pep  
Perfect Score: 3384  
Sequence: 1 MFLLATLYFALPLDLMSAE.....PVLMTLALALLSVLAETS 468

Scoring table:  
PAM 150  
Gap 11

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 08  
Listing first 45 summaries

## Database:

sptrembl5  
1:sp\_fungi 2:sp\_human 3:sp\_invertebrate 4:sp\_mammal  
5:sp\_mhc 6:sp\_organelle 7:sp\_phase 8:sp\_plant  
9:sp\_bacteria 10:sp\_rodent 11:sp\_virus 12:sp\_vertebrate  
13:sp\_unclassified

Statistics: Mean 46.587; Variance 79.510; scale 0.586

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3341	98.7	468	10	O35246	GDNF RECEPTOR ALPHA.	0.00e+00
2	3291	97.3	463	10	O35748	GDNFR-ALPHA/RRRI-DELTA	0.00e+00
3	3248	96.0	463	10	O35252	GDNF RECEPTOR BETA.	0.00e+00
4	3091	91.3	460	2	O15507	RET LIGAND 1.	0.00e+00
5	1562	46.2	464	10	O35977	GLIAL CELL LINE-DERIVE	0.00e+00
6	1556	46.0	464	2	O15316	GLIAL CELL LINE-DERIVE	0.00e+00
7	1556	46.0	464	2	O15328	RET LIGAND 2.	0.00e+00
8	696	20.6	397	10	O35118	GFRALPHA-3.	4.86e-11
9	696	20.6	397	10	O35325	GLIAL CELL LINE-DERIVE	4.86e-11
10	100	3.0	129	11	O97845	MATRIX PROTEIN P17 (FR	2.57e-01
11	100	3.0	130	11	O36786	MA-P17 (FRAGMENT).	2.57e-01
12	100	3.0	130	11	O36808	MA-P17 (FRAGMENT).	2.57e-01
13	100	3.0	133	11	O97725	MATRIX PROTEIN P17 (FR	2.57e-01
14	101	3.0	262	9	O50381	DNA PROTEIN (FRAGMENT	1.85e-01
15	103	3.0	399	9	O50790	ORIGIN OF REPLICATION	9.44e-02
16	100	3.0	486	11	O77804	GAG PROTEIN.	2.57e-01
17	103	3.0	492	11	O80624	GAG PROTEIN.	9.44e-02
18	98	2.9	130	11	O36772	MA-P17 (FRAGMENT).	4.95e-01
19	99	2.9	243	11	P89796	GAG PROTEIN (FRAGMENT)	3.57e-01
20	98	2.9	359	2	O15268	SKAP55 PROTEIN.	4.95e-01

ID	Score	Query Match	Length	DB	ID	Description	Pred. No.
21	98	2.9	1127	3	O94248	CODED FOR BY C. ELEGAN	4.95e-01
22	97	2.9	1217	9	O45440	MNGX.	6.85e-01
23	96	2.8	70	11	P89780	GAG PROTEIN (FRAGMENT)	9.44e-01
24	95	2.8	129	11	O36910	MA-P17 (FRAGMENT).	1.30e+00
25	96	2.8	130	11	O36773	MA-P17 (FRAGMENT).	9.44e-01
26	96	2.8	130	11	O76946	GAG PROTEIN (FRAGMENT)	9.44e-01
27	96	2.8	130	11	O78625	GAG PROTEIN (FRAGMENT)	9.44e-01
28	96	2.8	130	11	O76944	GAG PROTEIN (FRAGMENT)	9.44e-01
29	95	2.8	130	11	O75884	GAG PROTEIN (FRAGMENT)	1.30e+00
30	95	2.8	130	11	O75891	GAG PROTEIN (FRAGMENT)	1.30e+00
31	95	2.8	130	11	O75890	GAG PROTEIN (FRAGMENT)	1.30e+00
32	95	2.8	130	11	O75868	GAG PROTEIN (FRAGMENT)	1.30e+00
33	95	2.8	130	11	O75866	GAG PROTEIN (FRAGMENT)	1.30e+00
34	95	2.8	130	11	O75873	GAG PROTEIN (FRAGMENT)	1.30e+00
35	95	2.8	130	11	O75870	GAG PROTEIN (FRAGMENT)	1.30e+00
36	95	2.8	130	11	O75883	GAG PROTEIN (FRAGMENT)	1.30e+00
37	95	2.8	130	11	O75871	GAG PROTEIN (FRAGMENT)	1.30e+00
38	95	2.8	130	11	O75871	GAG PROTEIN (FRAGMENT)	1.30e+00
39	95	2.8	130	11	O75867	GAG PROTEIN (FRAGMENT)	1.30e+00
40	96	2.8	133	11	O97724	MATRIX PROTEIN P17 (FR	9.44e-01
41	95	2.8	242	11	O79070	GAG PROTEIN.	1.30e+00
42	96	2.8	304	11	O77701	GAG PROTEIN.	9.44e-01
43	96	2.8	492	11	O79351	GAG PROTEIN.	9.44e-01
44	95	2.8	492	11	O74452	GAG POLYPEPTIDE.	1.30e+00
45	95	2.8	623	8	O39204	BHLH PROTEIN.	1.30e+00

## ALIGNMENTS

RESULT 1. PRELIMINARY; PRT; 468 AA.  
ID O35246;  
AC O35246;  
DT 01-JAN-1998 (TREMBLREL. 05. CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05. LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05. LAST ANNOTATION UPDATE)  
DE GDNF RECEPTOR ALPHA.  
GN GDNF-ALPHA.  
OC MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C57; TISSUE=LIVER;  
RA DEV B.K.; WONG Y.W.; TOO H.P.;  
RL NEURORREPORT 9:0-0(0001).  
[2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C57; TISSUE=LIVER;  
RA DEV B.K.; WONG Y.W.; TOO H.P.;  
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF014117; G2624961; -;  
SQ SEQUENCE 468 AA: 51752 MW: AFDCE6A1 CRC32:

Query Match 98.7%; Score 3341; DB 10; Length 468;  
Best Local Similarity 97.6%; Pred. No. 0.00e+00;  
Matches 457; Conservative 10; Mismatches 1; Indels 0; Gaps 0;

DB	1	MFLLATLYFVLPDLDLMSAEVSGGDRDLCVKASDCLAEQSCSTRYRLRQCVAKETNF	60
OY	1	MFLLATLYFALPLDLMSAEVSGGDRDLCVKASDCLAEQSCSTRYRLRQCVAKETNF	60
DB	61	STLSGLEAKDECRSMALAKOKSLYNCRCRGMKEKNCRLRYWMSYOSLOGNDLLEDSR	120
OY	61	STLSGLEAKDECRSMALAKOKSLYNCRCRGMKEKNCRLRYWMSYOSLOGNDLLEDSR	120
DB	121	YEPVNSRLSDIRAPFISDVPOQVEHISKGNCCDAKACNLDTCKKRSAYITPCTT	180
OY	121	YEPVNSRLSDIRAPFISDVPOQVEHISKGNCCDAKACNLDTCKKRSAYITPCTT	180
DB	181	SNSNEVCNRKCKHALROFEDKVPKHSYGMFCSCPDVACTERRRQTIIVPCSYEERER	240
OY	181	SNSNEVCNRKCKHALROFEDKVPKHSYGMFCSCPDVACTERRRQTIIVPCSYEERER	240



```

Db 241 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 300
| | | | |
Oy 241 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 300
Db 301 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 360
| | | | |
Oy 301 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 360
Db 361 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDNDY 420
| | | | |
Oy 361 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDNDY 420
Db 421 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 468
| | | | |
Oy 421 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 468

RESULT 2
035748 PRELIMINARY; PRT; 463 AA.
01-JAN-1998 (TREMBLREL. 05, CREATED)
01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HANNOVER.
RA ZHONG J., ANNIES M., HEUMANN R.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AJ002072; E1154274; -
SQ SEQUENCE 463 AA; 51032 MW; 93277F91 CRC32;

Query Match 97.3%; Score 3291; DB 10; Length 463;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 463; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

Db 1 MFLATLYFALPLDLMLSAEYSGGDRDVCYKASPOCKLEQSCSTKYRTLRCQVAGKETNF 60
| | | | |
Oy 1 MFLATLYFALPLDLMLSAEYSGGDRDVCYKASPOCKLEQSCSTKYRTLRCQVAGKETNF 60
Db 61 SLTSGLEAKDECRSAMALOKSLYNCRCRGMKEKNCRLIYWSMYOSLOGNDLEDS 120
| | | | |
Oy 61 SLTSGLEAKDECRSAMALOKSLYNCRCRGMKEKNCRLIYWSMYOSLOGNDLEDS 120
Db 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 175
| | | | |
Oy 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 175
Db 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180
| | | | |
Oy 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180
Db 176 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 235
| | | | |
Oy 176 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 235
Db 181 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 240
| | | | |
Oy 181 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 240
Db 236 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 295
| | | | |
Oy 236 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 295
Db 241 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 300
| | | | |
Oy 241 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 300
Db 296 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 355
| | | | |
Oy 296 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 355
Db 301 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 360
| | | | |
Oy 301 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 360
Db 356 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDSD 415
| | | | |
Oy 356 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDSD 415
Db 361 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDSD 420
| | | | |
Oy 361 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDSD 420
Db 416 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 463
| | | | |
Oy 416 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 463
Db 421 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 468
| | | | |
Oy 421 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 468

RESULT 3

```

```

ID 035252 PRELIMINARY; PRT; 463 AA.
AC 035252;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE GDNFR-ALPHA/TRNR1-DELTA PROTEIN.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57;
RA DEY B.K., WONG Y.W., TOO H.P.;
RL NEUROREPORT 9:0-0(0001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-C57;
RA DEY B.K., WONG Y.W., TOO H.P.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF015172; G2624963; -
SQ SEQUENCE 463 AA; 51134 MW; 910EF17F CRC32;

Query Match 96.0%; Score 3248; DB 10; Length 463;
Best Local Similarity 96.6%; Pred. No. 0.00e+00;
Matches 452; Conservative 10; Mismatches 1; Indels 5; Gaps 1;

Db 1 MFLATLYFVPLDLMLSAEYSGGDRDVCYKASPOCKLEQSCSTKYRTLRCQVAGKETNF 60
| | | | |
Oy 1 MFLATLYFVPLDLMLSAEYSGGDRDVCYKASPOCKLEQSCSTKYRTLRCQVAGKETNF 60
Db 61 SLTSGLEAKDECRSAMALOKSLYNCRCRGMKEKNCRLIYWSMYOSLOGNDLEDS 120
| | | | |
Oy 61 SLTSGLEAKDECRSAMALOKSLYNCRCRGMKEKNCRLIYWSMYOSLOGNDLEDS 120
Db 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 175
| | | | |
Oy 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 175
Db 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180
| | | | |
Oy 121 YEPVNSRLSDIFRAVPFIS----VEHISKGNCLDAKACNLDTCCKYRSAYITPCTT 180
Db 176 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 235
| | | | |
Oy 176 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 235
Db 181 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 240
| | | | |
Oy 181 SMSNEVNCRRKCHALRQFDKVPKAKHSYGMFLCSCRIACTERRRQTIYVCSYEERER 240
Db 236 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 295
| | | | |
Oy 236 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 295
Db 241 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 300
| | | | |
Oy 241 PNCNLADSCNTNYICRSRLADFTNCPESRSVSNCKENYADCLAYSGLTGVTMPN 300
Db 296 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 355
| | | | |
Oy 296 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 355
Db 301 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 360
| | | | |
Oy 301 YIDSSLSVAAPWCDCSNGNDLEDCFLNFKDNTCLKNAIOAFGNGSDVTMMOPAPV 360
Db 356 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDNDY 415
| | | | |
Oy 356 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDNDY 415
Db 361 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDSD 420
| | | | |
Oy 361 OTTATTTTAFRIKPKPGAGSENEIPTHYLPPCANIOAKKLSNVSSTHCLSDSD 420
Db 416 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 463
| | | | |
Oy 416 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 463
Db 421 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 468
| | | | |
Oy 421 GRDGLAGASSHITTKSMAAPSCSLSPVMTALALSVSLAETS 468

RESULT 4
ID 015507 PRELIMINARY; PRT; 460 AA.
AC 015507;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE RET LIGAND 1.
GN RETL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.

```

RC TISSUE-KIDNEY;  
RA SANICOLA M., HESSION C.A., MORLEY D.S., CARMILLO P., EHRENFELS C.,  
RA VALUS L., ROBINSON S., JAWORSKI G., WEI H., TIZARD R., WHITTY A.,  
RA PEPIISKY R.B., CATE R.L.;  
RA PROC. NATL. ACAD. SCI. U.S.A. 94:6238-6243(1997).  
[2]  
RN  
RP TISSUE-KIDNEY;  
RC  
RA SANICOLA M., HESSION C.A., MORLEY D.S., CARMILLO P., EHRENFELS C.,  
RA VALUS L., ROBINSON S., JAWORSKI G., WEI H., TIZARD R., WHITTY A.,  
RA PEPIISKY R.B., CATE R.L.;  
RA SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
[3]  
RN  
RP TISSUE-KIDNEY;  
RC  
RA TISSUE-SUBSTANTIA NIGRA;  
RA HISHIKI T., KONDOH K., ICHIMITA S., NIMURA Y., SEKI N., OZAKI T.,  
RA SAKIYAMA S., TANAHASHI H., OHNUMA N., TANABE M., FUJIMURA S.,  
RA NAKAGAWA A.;  
RA SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
EMBL: U97144; G2282026;  
EMBL: U95847; G2459742;  
SEQUENCE 460 AA; 50838 MW; 022FECA CRC32;  
Query Match 91.3%; Score 3091; DB 2; Length 460;  
Best Local Similarity 91.8%; Pred. No. 0.00e+00;  
Matches 425; Conservative 27; Mismatches 5; Indels 6; Gaps 2;  
Db 1 MFLATYFALPLDLLLSAEVSGGDRDVCVAKASDQCKEQSCSTKRYTLROCVAGKETNF 60  
Oy 1 MFLATYFALPLDLLLSAEVSGGDRDVCVAKASDQCKEQSCSTKRYTLROCVAGKETNF 60  
Db 61 SLATGLEAKDCRSAMALOKSLYNCRCRGMKEKNCLEIYWSMOSLOGLNLLDSDP 120  
Oy 61 SLATGLEAKDCRSAMALOKSLYNCRCRGMKEKNCLEIYWSMOSLOGLNLLDSDP 120  
Db 121 YEPNRSRLSDIFRVPFIS-----VEHPIKGNCLDAKACNLDDICKRYSAITPCTT 175  
Oy 121 YEPNRSRLSDIFRVPFIS-----VEHPIKGNCLDAKACNLDDICKRYSAITPCTT 175  
Db 121 YEPNRSRLSDIFRVPFIS-----VEHPIKGNCLDAKACNLDDICKRYSAITPCTT 180  
Oy 121 YEPNRSRLSDIFRVPFIS-----VEHPIKGNCLDAKACNLDDICKRYSAITPCTT 180  
Db 176 SVSNDVCNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 235  
Oy 176 SVSNDVCNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 235  
Db 181 SMSNEVCNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 240  
Oy 181 SMSNEVCNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 240  
Db 236 PNCNLADSCNTYICRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 295  
Oy 236 PNCNLADSCNTYICRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 295  
Db 241 PNCNLADSCNTYICRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 300  
Oy 241 PNCNLADSCNTYICRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 300  
Db 296 YIDSSLSVAPMCDCSNGNLEDCIKFLNFKDNTCLKNAIOAFGNSDVTVMQAPAPV 355  
Oy 296 YIDSSLSVAPMCDCSNGNLEDCIKFLNFKDNTCLKNAIOAFGNSDVTVMQAPAPV 355  
Db 301 YIDSSLSVAPMCDCSNGNLEDCIKFLNFKDNTCLKNAIOAFGNSDVTVMQAPAPV 360  
Oy 301 YIDSSLSVAPMCDCSNGNLEDCIKFLNFKDNTCLKNAIOAFGNSDVTVMQAPAPV 360  
Db 356 QTTATATTTALRVKKNKPLGAPGSENEIPTHVLPCCANLQAKLSVNSGNTLCSNGY 415  
Oy 356 QTTATATTTALRVKKNKPLGAPGSENEIPTHVLPCCANLQAKLSVNSGNTLCSNGY 415  
Db 361 QTTATATTTARVKNKPLGAPGSENEIPTHVLPCCANLQAKLSVNSGNTLCSNGY 420  
Oy 361 QTTATATTTARVKNKPLGAPGSENEIPTHVLPCCANLQAKLSVNSGNTLCSNGY 420  
Db 416 EKREL-GASSHITTKSMAAPSCSLPLVLYVATLALSLT 457  
Oy 416 EKREL-GASSHITTKSMAAPSCSLPLVLYVATLALSLT 457  
Db 421 GKDLGAGASHITTKSMAAPSCSLPLVLYVATLALSLT 463  
Oy 421 GKDLGAGASHITTKSMAAPSCSLPLVLYVATLALSLT 463  
RESULT 5  
ID 035977 PRELIMINARY; PRT: 464 AA.  
AC 035977;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE GLIAL CELL LINE-DERIVED NEUROPROPHIC FACTOR RECEPTOR-BETA.  
GN GDNF-BETA OR RETL2  
OS RATUS NORVEGICUS (RAT)  
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA;  
OC EUTHERIA: RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.

RA TRUPP M., RAYNSCHER C., IBANEZ C.F.;  
RA SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN  
RP  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN/KIDNEY;  
RA SANICOLA M., HESSION C.A., MORLEY D.S., CARMILLO P., EHRENFELS C.,  
RA VALUS L., ROBINSON S., JAWORSKI G., WEI H., TIZARD R., WHITTY A.,  
RA PEPIISKY R.B., CATE R.L.;  
RA PROC. NATL. ACAD. SCI. U.S.A. 94:6238-6243(1997).  
[3]  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN/KIDNEY;  
RA SANICOLA M., HESSION C.A., MORLEY D.S., CARMILLO P., EHRENFELS C.,  
RA VALUS L., ROBINSON S., JAWORSKI G., WEI H., TIZARD R., WHITTY A.,  
RA PEPIISKY R.B., CATE R.L.;  
RA SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: AF005226; G2232252;  
DR EMBL: U97143; G2282024;  
SEQUENCE 464 AA; 51668 MW; B1A2BD11 CRC32;  
Query Match 46.2%; Score 1562; DB 10; Length 464;  
Best Local Similarity 47.9%; Pred. No. 0.00e+00;  
Matches 221; Conservative 100; Mismatches 126; Indels 14; Gaps 12;  
Db 12 FLDETSLASPSLSGSELHGMRPQVDYVANEICAEASNGSSRYTLROCVAGRDN- 70  
Oy 12 FLDETSLASPSLSGSELHGMRPQVDYVANEICAEASNGSSRYTLROCVAGRDN- 70  
Db 71 TM--LANK-ECQAALEVLDESPLYDCRCRGMKEKELQIYWSIMHGLTEGEFEYAS 136  
Oy 71 TM--LANK-ECQAALEVLDESPLYDCRCRGMKEKELQIYWSIMHGLTEGEFEYAS 136  
Db 61 SLTGLEAKDCRSAMALOKSLYNCRCRGMKEKNCLEIYWSMOSLOGLNLLDSDP 119  
Oy 61 SLTGLEAKDCRSAMALOKSLYNCRCRGMKEKNCLEIYWSMOSLOGLNLLDSDP 119  
Db 127 PYEYVRLSDIFRFLASIEGCTGDPVAVSTKSNHCLDAKACNLDDICKRYSAITPCTT 186  
Oy 127 PYEYVRLSDIFRFLASIEGCTGDPVAVSTKSNHCLDAKACNLDDICKRYSAITPCTT 186  
Db 120 PYEYVRLSDIFRFLASIEGCTGDPVAVSTKSNHCLDAKACNLDDICKRYSAITPCTT 179  
Oy 120 PYEYVRLSDIFRFLASIEGCTGDPVAVSTKSNHCLDAKACNLDDICKRYSAITPCTT 179  
Db 187 REISPTCRNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 246  
Oy 187 REISPTCRNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 246  
Db 180 TSM-NEVCNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 238  
Oy 180 TSM-NEVCNRKCHKALROFEDVPKHSYGLFSCSDIACIERRRQIIVPVSYEERK 238  
Db 247 EKPNCILRSICRDLHCRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 306  
Oy 247 EKPNCILRSICRDLHCRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 306  
Db 239 ERPNCLSDSCNTYICRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 298  
Oy 239 ERPNCLSDSCNTYICRSLADFTNCPESRSVSCLEKENVADCLAYSLIGVMTPN 298  
Db 307 PNYVDSNPTGIWSPMCNCGSGMEECEKFLRDFTEPNCLKNAIOAFGNSDVTVMQAPAPV 366  
Oy 307 PNYVDSNPTGIWSPMCNCGSGMEECEKFLRDFTEPNCLKNAIOAFGNSDVTVMQAPAPV 366  
Db 299 PNYVDS-S-LSVAPMCDCSNGNLEDCIKFLNFKDNTCLKNAIOAFGNSDVTVMQAP 356  
Oy 299 PNYVDS-S-LSVAPMCDCSNGNLEDCIKFLNFKDNTCLKNAIOAFGNSDVTVMQAP 356  
Db 367 GPSLPATQAPRV-E-KTPSLPDDLSGTS-IGTSVITCTSIQOGLKANSKELSCFT 423  
Oy 367 GPSLPATQAPRV-E-KTPSLPDDLSGTS-IGTSVITCTSIQOGLKANSKELSCFT 423  
Db 357 APVQTTATTTARVKNKPLGAPGSENEIPTHVLPCCANLQAKLSVNSGNTLCSNGY 416  
Oy 357 APVQTTATTTARVKNKPLGAPGSENEIPTHVLPCCANLQAKLSVNSGNTLCSNGY 416  
Db 424 ELTTNISPGSKYIKLNSGSRRLSALATLPLMLTLAL 464  
Oy 424 ELTTNISPGSKYIKLNSGSRRLSALATLPLMLTLAL 464  
Db 417 DSDFGKDLGAGASHITTKSMAAPSCSLPLVLYVATLALSLT 456  
Oy 417 DSDFGKDLGAGASHITTKSMAAPSCSLPLVLYVATLALSLT 456  
RESULT 6  
ID 015316 PRELIMINARY; PRT: 464 AA.  
AC 015316;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE GLIAL CELL LINE-DERIVED NEUROPROPHIC FACTOR RECEPTOR BETA.  
GN GDNF-BETA.  
OS HOMO SAPIENS (HUMAN)  
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA;  
OC EUTHERIA: PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-FETAL BRAIN;  
RA MARTIOVARA K., SUVANANTO P., HORELLI-KUITUNEN N., LINDAHL M.,  
RA MOSHNYAKOV M., AIRAKSINEN M.S., PALOTIE A., SARIOJA H., SAARMAN M.;



OY 236 EERERENCISLQSCKTNYICRSRLADFTFNCPESRSVSNCKENYADCLLAYSGLIGT 295  
 DB 296 AMPNPIISKVNTVALSCTCRGSGNLODECEQLEERSFSONPCLVEAIAA 344  
 OY 296 VMTPNYVSSLSVAPWCDCSNGNDLEDCLEKFLNFKMTCKLMAIOA 344  
 RESULT 9  
 ID 035325 PRELIMINARY; PRT; 397 AA.  
 AC 035325;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
 GN GLIAL CELL LINE-DERIVED NEUTROPHILIC FACTOR FAMILY RECEPTOR ALPHA-3.  
 OS GFRA-3.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 RN EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TRUPP M., RAYNSCHER C., IBANEZ C.F.;  
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF020305; G2429367; -  
 SQ SEQUENCE 397 AA; 44333 MW; F0C0C841 CRC32;  
 Query Match 20.68; Score 696; DB 10; Length 397;  
 Best Local Similarity 34.48; Pred. No. 4,86e-131;  
 Matches 120; Conservative 71; Mismatches 137; Indels 21; Gaps 16;  
 DB 11 LLMILLVLSLMLPLGNGSLATENFVNSCTQARKKCEANPCKAAYOHLSGCTSSLSR 70  
 OY 2 FLATLFLALPL-LDLMSAEVSGGDR-L-D-CYKASDQCLEQSCSTKRYLRQCVAGKET 58  
 DB 71 PLPLEBS-AMSAQDLAEAEQLRNSSLIDCRCHRMKHOATCLDIYTVHPARSLGDELD 129  
 OY 59 NFSLTSGLEAKDEKDCRSMEALQKOKSLYNCRCRGMKKEKCLRTIVSMYQSLQ-GNDLLE 117  
 DB 130 VSYE-E-DVYTSKPMK-MN-LSKL-NMLK--PPSDCLKRAMCTLHDKCDRLKRAYGEA 192  
 OY 118 DSEYEVENSLSLDIFRAVPEISDVFOVEHISGNCLDAKACNMDTCTCKYRSAYITP 177  
 DB 183 CS-GIR---CORHLCLQALRSFPEKAESAOGILLCPPEPADAGGERRNITAPSCAL 238  
 OY 178 CITSMSNEVCNRKCKHAKLQFQDKYPAKHSYGMLEFCSCR-DIACERRROTIVPVCSY 235  
 DB 239 PS-VTPNCIDLRSCFADPLCRSLMDFOHCHPMDI-LGTCAEQ-SRCLRAYLGLIGT 295  
 OY 236 EERERENCISLQSCKTNYICRSRLADFTFNCPESRSVSNCKENYADCLLAYSGLIGT 295  
 DB 296 AMPNPIISKVNTVALSCTCRGSGNLODECEQLEERSFSONPCLVEAIAA 344  
 OY 296 VMTPNYVSSLSVAPWCDCSNGNDLEDCLEKFLNFKMTCKLMAIOA 344  
 RESULT 10  
 ID 097845 PRELIMINARY; PRT; 129 AA.  
 AC 097845;  
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
 DE 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)  
 GN MATRIX PROTEIN P17 (FRAGMENT).  
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).  
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;  
 RN LENTIVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-BAB INT.  
 RA NARVA R., ROQUES P., COURBOTIN C., PARNETWATHIEU F., BOUSSIN F.,  
 RA ROANE A., MARC D., LASFARGUES G., DORMONT D.;  
 RL J. VIROL. 70:4474-4483(1996).  
 DR EMBL; X99948; E258593; -

KW MATRIX PROTEIN.  
 FT NON\_TER 129  
 SQ SEQUENCE 129 AA; 14550 MW; 542AB924 CRC32;  
 Query Match 3.0%; Score 100; DB 11; Length 129;  
 Best Local Similarity 41.9%; Pred. No. 2.57e-01;  
 Matches 18; Conservative 11; Mismatches 9; Indels 5; Gaps 5;  
 DB 28 KYR-LKHIYASRELEFALNPGLESAGCCQQLMEQL-QSTL 68  
 OY 45 KYRTLQCV-AGKETN-FSLTSGL-EAKDECRSMEALQKSL 84  
 RESULT 11  
 ID 036786 PRELIMINARY; PRT; 130 AA.  
 AC 036786;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
 GN MA-P17 (FRAGMENT).  
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).  
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;  
 RN LENTIVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LEIGH BROWN A.J., LOBIDEL D., MADE C.M., REBUS S., PHILLIPS N.,  
 RA BRETTLE R.P., FRANCE A.J., LEEN C.S., MCENAMIN J., MCILLAN A.,  
 RA MAW R.D., MULCAHY F., ROBERTSON J.R., SANKAR K.N., SCOTT G., WYLD R.,  
 RA PEUTHERER J.F.;  
 RL VIROLOGY 235:166-177(1997).  
 DR EMBL; AF014183; G2406708; -  
 FT NON\_TER 130  
 SQ SEQUENCE 130 AA; 14584 MW; 428D9E75 CRC32;  
 Query Match 3.0%; Score 100; DB 11; Length 130;  
 Best Local Similarity 41.9%; Pred. No. 2.57e-01;  
 Matches 18; Conservative 11; Mismatches 9; Indels 5; Gaps 5;  
 DB 5 KYR-LKHIYASRELEFALNPGLESAGCCQQLMEQL-QPSL 45  
 OY 45 KYRTLQCV-AGKETN-FSLTSGL-EAKDECRSMEALQKSL 84  
 RESULT 12  
 ID 036808 PRELIMINARY; PRT; 130 AA.  
 AC 036808;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
 GN MA-P17 (FRAGMENT).  
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).  
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;  
 RN LENTIVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LEIGH BROWN A.J., LOBIDEL D., MADE C.M., REBUS S., PHILLIPS N.,  
 RA BRETTLE R.P., FRANCE A.J., LEEN C.S., MCENAMIN J., MCILLAN A.,  
 RA MAW R.D., MULCAHY F., ROBERTSON J.R., SANKAR K.N., SCOTT G., WYLD R.,  
 RA PEUTHERER J.F.;  
 RL VIROLOGY 235:166-177(1997).  
 DR EMBL; AF014205; G2406752; -  
 FT NON\_TER 130  
 SQ SEQUENCE 130 AA; 14527 MW; 7325723C CRC32;  
 Query Match 3.0%; Score 100; DB 11; Length 130;  
 Best Local Similarity 39.5%; Pred. No. 2.57e-01;  
 Matches 17; Conservative 12; Mismatches 9; Indels 5; Gaps 5;  
 DB 5 KYR-LKHIYASRELEFALNPGLETSEGCROIMEQL-QSAL 45

OY 45 KYRTLRQCV-AGKETN-FSLTSGT-EAKDECRSMEALKOKSL 84

RESULT 13  
ID 097725 PRELIMINARY; PRT; 133 AA.

AC 097725;  
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)  
DE MAIRIX PROTEIN P17 (FRAGMENT).  
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).  
OC VIRIDAE: SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;  
OC LENTIVIRIDAE.

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PAL-M:  
RA NAWA R., ROQUES P., COURPOTIN C., PARNETMATHIEU F., BOUSSIN F.,  
J. VIROL. 70:4474-4483(1996).  
EMBL: 279562; E261908; -.

KM MATRIX PROTEIN.  
FT NON\_TER 133  
SQ SEQUENCE 133 AA; 14724 MW; E22E819C CRC32;

Query Match 3.0%; Score 100; DB 11; Length 133;  
Best Local Similarity 44.28; Pred. No. 2,576-01;  
Matches 19; Conservative 9; Mismatches 10; Indels 5; Gaps 5;

Db 28 KYR-LKHIVASRELEFALNGLLEASGCRQLGOL-QPSL 68  
OY 45 KYRTLRQCV-AGKETN-FSLTSGT-EAKDECRSMEALKOKSL 84

RESULT 14  
ID 050381 PRELIMINARY; PRT; 262 AA.

AC 050381;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)  
DE DNAN PROTEIN (FRAGMENT).  
GN DNAN.  
OS MYCOBACTERIUM SMEGMATIS.  
OC PROKARYOTA: FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-2-WC155;  
MEDLINE: 96062237.  
RAJAGOPALAN M., QIN M.H., NASH D.R., MADIRAJU M.V.V.S.;  
J. BACTERIOL. 177:6527-6535(1995).  
EMBL: U17833; G1079500; -.

DR NON\_TER 262  
SQ SEQUENCE 262 AA; 27269 MW; E0114FC9 CRC32;

Query Match 3.0%; Score 101; DB 9; Length 262;  
Best Local Similarity 30.6%; Pred. No. 1,856-01;  
Matches 22; Conservative 14; Mismatches 33; Indels 3; Gaps 3;

Db 172 ESYVLAATDRFLAVRELITWTAGDEAVLVPAKTL-AEAKAGTGDGNOVHLGSGA 230  
OY 361 OTTATTTTAFVKNKPLGPAGSENEIPTHYLPFCANLQAOKLSNVSQS-THLCT-SDS 418

Db 231 SVGKGLGIRS 242  
OY 419 DFGKDLGAGSS 430

RESULT 15  
ID 050790 PRELIMINARY; PRT; 399 AA.

AC 050790;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)

DE ORIGIN OF REPLICATION AND GENES RNPA, RPMH, DNAA,  
DE DNAN, RECF.

GN DNAN.  
OS MYCOBACTERIUM TUBERCULOSIS.  
OC PROKARYOTA: FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV.  
RA SALAZAR L., FSIHI H., DE ROSSI E., RICCARDI G., RIOS C.,  
RA COLE S.T., TAKIF H.E.;  
RL MICROBIOLOGY 20:283-293(1996).  
EMBL: X92504; E208997; -.

DR EMBL: X92504; E208997; -.  
SQ SEQUENCE 399 AA; 41773 MW; 640BC587 CRC32;

Query Match 3.0%; Score 103; DB 9; Length 399;  
Best Local Similarity 27.78; Pred. No. 9,446-02;  
Matches 28; Conservative 27; Mismatches 42; Indels 4; Gaps 4;

Db 178 AYDSRLAVRELEMSASSPDIEAVLVPAKTL-AEAKAGIGSDVRLSTGPGVGKDL 236  
OY 368 TTARFVKNKPLGPAGSENEIPTHYLPFCANLQAOKLSNVSQS-THLCT-SDSDFGKDL 425

Db 237 LGISGN-GKRSTRLDAEPKFRQLLPTEHTAVATMDVAE 276  
OY 426 AGASSHITTKSMAAPSCSLSLPVLMTALALSLVSLAE 466

Search completed: Tue Jun 23 18:32:04 1998  
Job time : 133 secs.